

## Repositoris digitals

Disseny i implementació per a biblioteques, arxius i museus

Ciro Llueca

Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya  
18 i 25 de gener; 1 i 8 de febrer de 2013



*Agraïment especial de l'instructor a **Sandra Reoyo**, responsable de repositoris del Consorci de Biblioteques Universitàries de Catalunya, i coautora de les idees i materials que serveixen de base al curs.*

*Agraïment del formador a les persones que han participat en la preparació del curs:*

*Isidro Aguillo (CSIC); Joan Boadas (AMGi); Cristina Azorín (UAB); Tomàs Baiget (E-LIS); Isabel Bernal (CSIC); Manuel Bravo (BVMC); David Iglesias (AMGi); Francesca Navarro (BC); Patricia Russo (COBDC); Eugènia Serra (BC); Imma Subirats (E-LIS); Meritxell Trullàs (BdG); Ricard de la Vega (CESCA).*

## Objectiu

*Conèixer i analitzar els aspectes relacionats amb el disseny, implementació i avaluació de repositoris digitals des d'una perspectiva de compliment de l'estratègia de l'organització en la qual s'integra aquest servei.*

*La prioritat és aprendre a crear i gestionar repositoris digitals sostenibles.*

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Programa

1. **Què són els repositoris digitals i per a què serveixen**
  - 1.1 Cultura *Open Access*: de la *serial crisis* a la resposta de la comunitat
  - 1.2 Tipus de repositoris: institucionals, cooperatius, temàtics, recol·lectors
2. **Disseny i implementació d'un nou repositori**
  - 2.1 Cobertura estratègica i legal: rere la coherència amb la institució i la llei
  - 2.2 Maquinari i programari: sistemes propietaris, lliures i a mida
  - 2.3 Planificació dels processos de treball
3. **Difusió i avaluació**
  - 3.1 Publicació i promoció: objectiu "repositoris sostenibles"
  - 3.2 Avaluació: visibilitat, interoperabilitat i polítiques públiques
4. **Preservació digital**
  - 4.1 Estratègia: definint una política per al futur
  - 4.2 Tàctica: a l'espera de la solució global, algunes recomanacions

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Formador

**Ciro Llueca**, [cllueca@gmail.com](mailto:cllueca@gmail.com)

- Llicenciat en Documentació
- Coordinador de PADICAT a la Biblioteca de Catalunya
- Professor del Grau d'Informació i Documentació de la UB i del Màster en Documentació Digital de la UPF
- Editor per a Espanya i coordinador editorial per a Europa del repositori E-LIS

**Materials realitzats per **Ciro Llueca & Sandra Reoyo****

Disponibles a E-LIS i Slideshare.

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## 1. Què són els repositoris digitals i per a què serveixen

### 1.1 Cultura *Open Access*

#### Què és l'*Open Acces* / Accés lliure?

- Moviment mundial posat en marxa per les institucions que financen la recerca amb la finalitat de donar accés a la producció científica a través de la xarxa de forma gratuïta
- Es basa en la idea que els resultats de la recerca finançada amb diners públics han de ser d'accés públic

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

#### Què és l'*Open Acces* / Accés lliure?



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

Font: www.plos.org

#### Causes del moviment *Open Access*

- Desproporcionada pujada de preus de les revistes científiques (*serial crisis*, 80s)
- Control dels drets de *copyright* sobre els treballs publicats
- Imposicions de les editorials en polítiques d'accés i distribució (*big deals*)
- Avenços de les noves tecnologies i Internet
- Resposta de la comunitat científica i de la societat davant dels abusos editorials
- Llibertat intrínseca d'internet per al tràfic i compartiment de la informació

(Melero, 2005)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

#### *Serial crisis* de les revistes (1)

- Creixement de la producció i publicació científica des dels anys 60, de les biblioteques i de les seves col·leccions
  - Llei del creixement exponencial 10-15 anys (Price, 1963)
  - Es publica el doble en 20 anys (1980-2000) (Cox, 2000)
- Puja el preu de les publicacions
  - De les revistes, un 207% (1986-1999); el pressupost per a revistes creix un 170%
  - De les monografies, un 65% (1986-1999); el pressupost per a monografies creix un 34% (Association of Research Libraries, 2000)



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Serial crisis de les revistes (2)

- Les editorials passen de ser universitats i centres d'investigació a monopolis econòmics i culturals de publicació-distribució
  - 31 països van concentrar el 98% dels articles altament citats, mentre 162 països només van contribuir el 2% del total (King, 2004)
  - Dels 1.222 científics més citats, el 66% treballen en institucions dels USA, el 20,5% al Regne Unit, el 5,1% a Alemanya (King, 2004)
  - "Los precios de las editoriales científicas hacen parecer a Rupert Murdoch un socialista" (Monbiot, 2011)
    - 1 setmana de l'accés total al diari britànic "The Times" = 9,6 USD (thetimes.com)
    - 1 accés a 1 article d'Elsevier = 31,50 USD ; Springer = 34,95 USD ; Wiley-Blackwell = 42 USD (monbiot.com, 2011)

I want to purchase this article

Price: USD 32

In order to purchase this article you must first purchase a copy



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Posició de força [i abús] de les editorials

- Preus
- Monopoli cultural i empresarial
- Control dels drets de *copyright*
- Imposicions en polítiques d'accés i distribució
- Nul·la visibilitat de la literatura gris



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

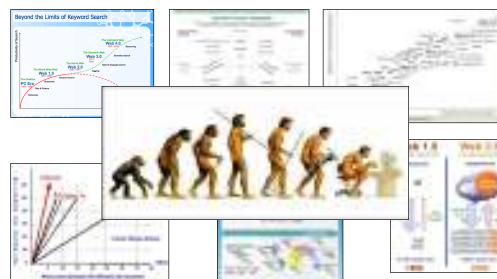
- "La Revista Iberoamericana de Automática es por segundo año consecutivo la única revista española de Informática en el *Journal Citation Reports*" (ucm.es 2011)
- Localitza, a la web de la revista (<http://www.revista-riai.org/>), les condicions per autors relatives al *copyright* de la publicació i avalua què pot fer un autor amb el seu article, si opta per presentar-lo allà.

10' treball individual +  
5' debat conjunt



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Escenari tecnològic



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

Free Online Scholarship Movement

## La resposta de la comunitat: 3B (1)

### Declaració de Budapest (2001)

- "Hacemos un llamado a todas las instituciones e individuos interesados para que ayuden a incrementar al acceso abierto al resto de este tipo de literatura y retiren las barreras, en especial las barreras de precio que se interponen en este camino"
- "Por "acceso abierto" a esta literatura queremos decir su disponibilidad gratuita en Internet público, permitiendo a cualquier usuario leer, descargar, copiar, distribuir, imprimir, buscar o usarlos con cualquier propósito legal (...) el único rol del *copyright* (...) deberá ser dar a los autores el control sobre la integridad de sus trabajos y el derecho de ser adecuadamente reconocidos y citados"
- Recomendamos dos estrategias: Auto Archivo y Publicaciones periódicas de acceso abierto

<http://www.soros.org/openaccess/translations/spanish-translation>

Budapest Open Access Initiative

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## La resposta de la comunitat: 3B (2)

### Declaració de Bethesda (2003)

- Una publicació serà d'accés obert si "una versió completa de la obra y todos los materiales suplementarios (...) se deposita de forma inmediata a la publicación inicial en al menos un repositorio en línea apoyado por una institución académica..."
- "La investigación científica es un proceso interdependiente dónde cada experimento se nutre de los resultados de otros"

[http://ictlogy.net/articles/bethesda\\_es.html](http://ictlogy.net/articles/bethesda_es.html)

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## La resposta de la comunitat: 3B (3)

### Declaració de Berlin (2003)

- “Nuestra misión de diseminar el conocimiento será incompleta si la información no es puesta a disposición de la sociedad de manera expedita y amplia”
- “Se requiere idealmente el compromiso activo de todos y cada uno de quienes producen conocimiento científico y mantienen el patrimonio cultural”
- “Nuestras organizaciones tienen el propósito de encontrar soluciones que ayuden a un mayor desarrollo de los marcos legales y financieros existentes, con el fin de facilitar óptimo uso y acceso”

[http://www.geotropico.org/1\\_2\\_Documentos\\_Berlin.html](http://www.geotropico.org/1_2_Documentos_Berlin.html)

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## La resposta de la comunitat (4)

- Protocol d'intercanvi de registres OAI-PMH (Open Archives Initiative-Protocol Metadata Harvesting, 2002- )
- ROARMAP (Registry of Open Access Repository Material Archiving Policies, 2008- ) i MELIBEA (Directorio y estimador de políticas en favor del acceso abierto a la producción científica, 2010- )
  - 338 polítiques institucionals inscrites a ROARMAP, 247 a MELIBEA
- OpenAIRE (Open Access Infrastructure for Research in Europe, Comissió Europea, 2010)
  - “Todos los ciudadanos de la UE tienen el derecho a acceder a los conocimientos generados gracias a los fondos públicos y a sacar ventaja de los mismos”
- Declaració de La Alhambra (Sud d'Europa, 2010)
  - La importància d'implementar polítiques (institucionals, de consorci, nacionals) per fomentar l'OA: dipòsit obligatori, transició de les revistes a OA, incentius per complir OA, acceptar material dipositat per a l'avaluació de la investigació, etc.
- Open Access Week (2010- )
  - Learn. Share. Advance



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

- Localitza en el registre de ROARMAP <http://roarmap.eprints.org/>, o MELIBEA <http://www.accesoabierto.net/politicas/>, el document d'una institució i identifica les característiques bàsiques del mandat

10' treball individual +  
10' debat conjunt



## La resposta de la comunitat (5)

### A Espanya

#### • Ley 14/2011 de la Ciencia, la Tecnología y la Innovación

- “El personal de investigación cuya actividad investigadora esté financiada mayoritariamente con fondos de los Presupuestos Generales del Estado hará pública una versión digital de la versión (...) tan pronto como resulte posible, pero no más tarde de doce meses después de la fecha oficial de publicación”
- “La versión electrónica se hará pública en repositorios de acceso abierto reconocidos en el campo de conocimiento en el que se ha desarrollado la investigación, o en repositorios institucionales de acceso abierto”

#### • Real decreto 99/2001 de Enseñanzas de doctorado

- Art. 14. Evaluación y defensa de la tesis doctoral: “5. Una vez aprobada la tesis doctoral, la universidad se ocupará de su archivo en formato electrónico abierto en un repositorio institucional...”

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## La resposta de la comunitat (6)

### A Catalunya

- Aprobació del document marc del mandat d'accés obert de la recerca de les universitats de Catalunya elaborat pel Consorci de Biblioteques Universitàries de Catalunya (Consell Interuniversitari de Catalunya, 2010)
- Polítiques institucionals d'accés obert UPC, UOC, UPF, UB, UdG, UAB i UdL
- Acord de la Junta del Consell Interuniversitari de Catalunya sobre els estudis de doctorat i el procediment per a la creació de les escoles de doctorat (Consell Interuniversitari de Catalunya, 2011)

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## La resposta de la comunitat (7)

- Existeixen dues estratègies recomanades per a la implementació de l'OA:

- La ruta verda (*green road*), impulsa el dipòsit de recursos digitals en repositoris digitals
  - El directori **OpenDOAR** (Directory of Open Access Repositories, 2006- ) conté més de 2.200 repositoris digitals, la majoria universitaris. D'aquests 94 són espanyols
- La ruta dorada (*gold road*), promou la publicació en revistes *open access*
  - **DOAJ** (Directory of Open Access Journals, 2001- ) inclou més de 8.100 revistes OA
- Aquestes dues estratègies han donat lloc a models híbrids (pagament per subscripció + accés obert a determinats continguts; l'autor/institució paga per publicar; etc.)

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## OpenAccess.es

- 2000: revistes OA, grups de treball, articles
- 2001: TDX (Tesis Doctorals en Xarxa), Consorci Biblioteques Universitàries de Catalunya (CBUC)
- 2003: Canàries firma Declaració de Berlin. E-LIS (E-prints in Library and Information Science)
- 2006: 12 repositoris, Hispana, OS-Repositorios
- 2008: Recolecta, Dulcinea, Melibea, Buscarepositorios Ranking Web of Repositories (CSIC)
- 2011: Ley 14/2011, Real Decreto 99/2011
- 2012: 94 repositoris (4,3% prod. mundial)

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

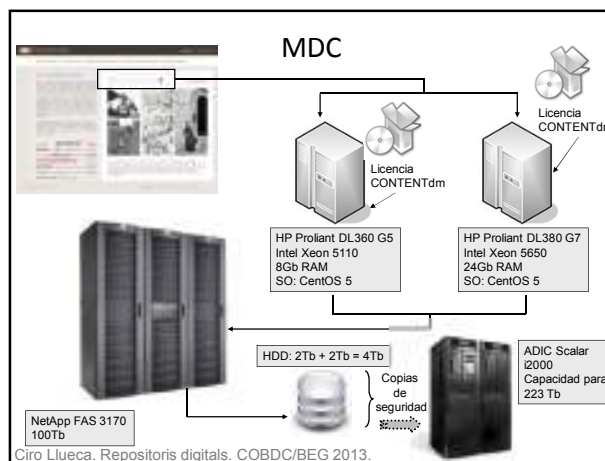
## Pràctica

- Consulta el directori OpenDOAR (Directory of Open Access Repositories, 2006- )
- Conté més de 2.200 repositoris digitals, la majoria universitaris. D'aquests 94 són espanyols
- Localitza un repositori d'una temàtica del teu interès, familiaritzat amb el seu sistema i prepara una explicació concisa per als companys
- Coneixes algun repositori que no estigui a OpenDOAR? (Opció Suggest)

10' treball individual + 5' presentació



## 1.2 Tipus de repositoris



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Què són i per a què serveixen (1)

- “Un conjunto de servicios prestados por una institución a su comunidad para recopilar, gestionar, difundir y preservar su producción digital a través de una colección organizada, de acceso abierto e interoperable” (Barrueco et al., 2010)
- “Un repositorio, depósito o archivo es un sitio centralizado donde se almacena y mantiene información digital, habitualmente bases de datos o archivos informáticos” (Wikipedia, 2011)
- “Un repositorio institucional es una base de datos compuesta de un grupo de servicios destinados a capturar, almacenar, ordenar, preservar y redistribuir la documentación académica de la universidad en formato digital” (Barton i Waters, 2005)

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Què són i per a què serveixen (2)

- Un repositori serveix per publicar a Internet, normalment amb accés lliure, una base de dades que inclou els documents digitals o digitalitzats que formen els fons i col·leccions relacionades amb una unitat d'informació
- Els repositoris representen l'espai natural de creixement dels serveis d'informació: els processos de catalogació van donar lloc als catàlegs automatitzats, aquests als catàlegs en línia, i aquests als repositoris

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

### Què són i per a què serveixen (3)

- Un repositori que albergui la producció de la seva institució aconsegueix més visibilitat, difusió i prestigi dels seus autors; d'aquesta producció documental i científica; i dels seus centres d'investigació i difusió
- La visibilitat clàssica de la producció investigadora s'obre a l'entorn internacional, amb el que s'assegura més audiència i millor impacte científic

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

### Què són i per a què serveixen (4)

- A més, contribueixen al moviment mundial d'accés obert gràcies a que dipositen la seva producció acadèmica de forma gratuïta
- I possibilita una millor preservació dels seus fons, donat que es promou la digitalització retrospectiva de les seves col·leccions, i l'accés a la documentació original queda restringida a casos concrets

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

### Repositori institucional

- Un repositori institucional és el conjunt de serveis que una universitat ofereix als membres de la seva comunitat per a la gestió i difusió de material digital creat per la institució i els seus membres (Lynch, 2003)
- Crea pels serveis d'informació d'institucions acadèmiques per assegurar l'accés a la producció científica generada pels seus membres en el compliment de les seves funcions
- Normalment inclouen tesis, articles de revista, *pre* i *post-prints*, treballs de final de carrera, *working papers*, ponències i comunicacions de congressos, memòries, col·leccions especials, etc.

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

### ACCEDA



<http://acceda.ulpgc.es/>

### ACCEDA

- La Universidad de Las Palmas de Gran Canaria és una universitat fundada al 1989. 1.500 professors, 55 titulacions de grau i 37 programes de doctorat per a més de 24.000 estudiants
- El repositori Acceda (2009-) facilita el dipòsit i accés a tesis, treballs de grau i postgrau, literatura gris i informes de departament, etc., a la comunitat d'estudiants i professors de la seva universitat
- DSpace, 5.000 registres
- S'incentiva l'autoarxiu (video d'ajuda, reclam en la portada) i els formats audiovisuals
- El dipòsit es complementa amb un blog dedicat a l'accés obert, <http://oaulpgc.wordpress.com/>

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

### Pràctica

- ¿Tothom pot publicar articles científics a ACCEDA?
- Localitza al repositori ACCEDA (<http://acceda.ulpgc.es/>):
  - Una emissió de video sobre la pobresa, a càrrec de Juan Torres.
  - Documents sobre volcans que afecten la illa de El Hierro

10' treball individual



## Repositori cooperatiu

- Creats entre diverses institucions relacionades entre elles per algun motiu (per ex.: per formar part del mateix territori) amb la finalitat de servir de repositori institucional a aquelles institucions que no puguin o no vulguin mantenir un repositori propi
- Tenen com a finalitat agrupar la documentació per a augmentar la seva visibilitat i la de la investigació produïda en el seu àmbit d'actuació; i obtenir serveis de valor afegit que difícilment es poden obtenir a nivell individual

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## RECERCAT



<http://www.recercat.cat>

## RECERCAT

- Repositori cooperatiu de documents de recerca de les universitats i centres d'investigació de Catalunya. Els seus objectius són:
  - Augmentar la visibilitat dels documents, dels autors i de les institucions que els emparen i, en general, de la investigació produïda a Catalunya;
  - Facilitar l'edició a través de l'entrada auto-alimentada dels documents;
  - Afegir valor als documents a través de citacions normalitzades, estadístiques de consulta, adreces permanents i mecanismes de preservació.
- El sistema inclou registres dipositats directament en el repositori, o bé dipositats en altres repositoris associats, sempre que compleixin metadades estàndards
- DSpace, més de 42.000 documents

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

- Localitza a RECERCAT, <http://www.recercat.cat>:
  - Un document sobre el perfil del gestor cultural a Espanya de L. Bernal. Està dipositat a RECERCAT?
  - Quants documents hi ha de l'Institut Barcelona d'Estudis Internacionals?

10' treball individual



## Repositori temàtic

- Creats amb la finalitat d'agrupar els documents d'investigació sobre una temàtica concreta
- Habitualment funcionen de manera suprainstitucional, a partir de les aportacions que realitzen els autors en servei d'auto-arxiu
- Gràcies a la seva especificitat temàtica i a les eines d'alerta que normalment ofereixen, són repositoris molt útils per als investigadors

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## E-LIS



<http://eprints.rclis.org/>

## E-LIS (2)

- **E-LIS** (E-prints in Library and Information Science, <http://eprints.rclis.org/>) fou creat el 2003 com repositori temàtic de l'àmbit de la Informació i la Documentació, amb l'objectiu de donar visibilitat als treballs científics d'aquestes matèries.
- És un projecte sense ànim de lucre finançat per diverses institucions, s'alimenta per mitjà de l'autoarxiu, i està mantingut per voluntaris.
- Compta amb un equip d'editors de 44 països i el suport a 22 llengües.
- **Imma Subirats** és la seva cofundadora i directora executiva: "Vull que E-LIS sigui una eina útil per a tots els bibliotecaris que volen millorar".

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



## E-LIS (3)

- Eprints (Dspace fins gen. 2013), 13.600 registres
- Creixement 853 documents anuals (2011).
- El depòsit (previ registre) i la consulta són gratuïts, inclou tot tipus de documentació científica escrita.
- Utilitza per a la descripció de recursos un tesaure propi: *JITA Classification Schema of Library and Information Science* (<http://eprints.rclis.org/cms/jita/>).

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



## Portal de repositoris

- En un nivell supracorporatiu es troben els portals que actuen de paraigües de diversos repositoris
- També anomenats recol·lectors / *harvesters* / agregadors
- Per exemple:
  - Europeana <http://www.europeana.eu>
  - DRIVER <http://www.driver-repository.eu>
  - SciELO <http://www.scielo.org/php/index.php>
  - OAster <http://oaister.worldcat.org/>

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## OAster



<http://oaister.worldcat.org/?&lang=es>

## OAster

- Catàleg col·lectiu de recursos d'accés obert creat a partir de les col·leccions de lliure accés en tot el món a través del protocol per a la recollida de metadades OAI-PMH
- OAster redirigeix els seus registres al repositori original que els conté
- 25 milions de registres, més de 1.100 col·laboradors

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

- Localitza al portal OAster (<http://oaister.worldcat.org/>) un mapa antic, de qualsevol ciutat europea:
  - Comprova les funcions de cerca avançada del catàleg, així com les de refinat de cerca.
  - Accedeix al document dipositat i comprova les funcions de consulta i edició del document.

10' treball individual





### 1.3 Estudi de cas: L'Arxiu Municipal de Girona

#### Arxiu Municipal de Girona (1)



<http://www.girona.cat/sqdap/cat/index.php>

#### Arxiu Municipal de Girona (2)

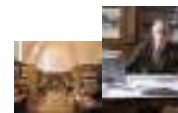
- L'Arxiu Municipal de Girona (AMG), inclou una secció d'Arxiu Administratiu i una altra d'Arxiu Històric (s. XIV- )
- És l'organ responsable d'assumir i gestionar la documentació municipal



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

#### Arxiu Municipal de Girona (3)

- Pressupost d'1 M€ euros al 2011
- Treballen 19 persones (15+4 autònoms). Depèn de l'ajuntament
- 40 TB de dipòsit digital
- El dirigeix Joan Boadas: "Sóc arxiver perquè m'interessa el futur"
- Al 1990 es va dissenyar una política rigurosa per donar accessibilitat a la documentació de l'AMG: el fons fotogràfic, el de premsa i el de documentació històrica de la ciutat
- Es produeix un canvi de filosofia que s'anticipa al seu temps:
  - De la custòdia i preservació de la documentació analògica;
  - A l'accés a través de la tecnologia i la especialització del centre
- 65.000 usuaris a l'any visiten el portal web



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

#### Arxiu Municipal de Girona (4)

- Fons d'imatges
  - Al 1982 es forma el fons d'imatges antigues de l'Ajuntament de Girona, a partir de col·leccions privades donades al municipi
  - Al 1990 es produeix (encàrrec informàtic extern) un programari a mida per a la gestió del fons fotogràfic: APCimatge. Al 1996 es comercialitza el programari, 15 llicències. Al 2005 canvi de programari
  - Es dissenya un sistema propi de metadades i quadre de classificació, basat en estàndards
  - Al 1997 es crea el CRDI (Centre de Recerca i Difusió de la Imatge), s'especialitza al personal responsable
  - A partir del càlcul de costos en el tractament de les fotografies (1,7 € x foto descrita, digitalitzada i instal·lada), es duu a terme una política agressiva de convenis de cooperació amb fotògrafs i propietaris locals que busca assumir fons privats no per compra, sinó per acords d'explotació (50/50%) de les imatges, un cop digitalitzades i incorporades a l'AMG
  - Es digitalitza cada imatge per a la seva preservació (màster) i difusió
  - El model s'exporta a fons videogràfics
  - 3.800.000 fotografies, 1.369 pel·lícules, 8.000 hores de video, 1.609 hores d'àudio



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

#### Arxiu Municipal de Girona (5)

- Fons de premsa
  - Al 1990 es dissenya l'objectiu de crear un portal d'informació local, que inclogui la documentació publicada en la premsa local, les emissores locals de televisió i ràdio, per a complementar l'arxiu fotogràfic
  - La documentació és extremadament fràgil però és clau per a la investigació de la història contemporània
  - Programari Pandora, 1,6 M pàgines de premsa

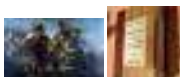


Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Arxiu Municipal de Girona (6)

- Fons documental històric
  - Al 1990 s'inicia la digitalització retrospectiva del fons administratiu històric des d'una perspectiva d'ordre cronològic invers, per a ser més útil als processos administratius de l'ajuntament
  - Es digitalitzen els llibres d'actes de 1601 a 1911
  - Programari Pandora, 786.616 pàgines digitalitzades
- Es formen "centres temàtics d'interés" dedicats a fets històrics rellevants. Paquets documentals que són la porta d'accés a investigadors i ciutadans: Guerra contra França (s. XIX), Judaisme, etc.
- Es crea un servei d'efemerides

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



## Pràctica

- Localitza i accedeix al document a "Premsa digitalitzada" de la web de l'AMG: (<http://www.girona.cat/sgdap/cat/premsa.php>):
  - Notícies de les inundacions de 1887.
- Accedeix a la secció d'efemerides a la documentació relativa a una data de la teva elecció: (<http://www.girona.cat/sgdap/cat/efemerides.php>)
- Accedeix al portal de "Patrimoni Audiovisual", fruit de la recerca de l'AMG.

10' treball individual



## 2. Disseny i implementació d'un nou repositori

### 2.1 Cobertura estratègica i legal

### Cobertura estratègica (1)

- Les unitats d'informació han de conèixer (i participar, si és possible) en la planificació estratègica de la institució
- Iniciar un procés de disseny i implementació d'un repositori digital ha de partir necessàriament d'un pla que sigui coherent amb l'entitat que finançarà i impulsarà la nova eina

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



### Cobertura estratègica (2)

- Sumar esforços entre àrees, departaments, seccions...
- Projectes (co)finançables i que comptin amb la complicitat i implicació dels directius
- Evitar la duplictat d'esforços

Per a què volem crear el repositori?  
És una necessitat real?

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



### Cobertura estratègica (3)

**Misión**

La misión de la Biblioteca Nacional es recoger el patrimonio bibliográfico y documental español, garantizar su integridad y facilitar el acceso a toda la ciudadanía y a las generaciones futuras.

**PLAN ESTRATÉGICO**  
2012-2014

**4. Reforzar las funciones institucionales de la Biblioteca Nacional**

- 4.1. Garantizar el acceso gratuito de usuarios no registrados al SIGB
- 4.2. Incremento de las colecciones y servicios digitales
- 4.3. Desarrollo del sistema de procesamiento digital
- 4.4. Mejora de la accesibilidad en línea y presencial
- 4.5. Fomento de los programas de investigación, tanto propios como en colaboración con otras instituciones
- 4.6. Desarrollo de programas profesionales
- 4.7. Mejora de los programas y herramientas de difusión

Ciro Lluéca. Repositoris digitals. COBDC/BEG 2013.

### Cobertura estratègica (4)

- Objectius marc del repositori
- Descripció dels recursos necessaris:
  - Humans
  - Infraestructura
  - Col·leccions (prioritzar les que tenen més impacte)
    - Documents nascuts digitals
    - Digitalització retrospectiva
- Dimensions de futur (creixement en docs., usos...)
  - Formats acceptats
  - Metadades



Ciro Lluéca. Repositoris digitals. COBDC/BEG 2013.

### Cobertura estratègica (5)

- Què opinen els investigadors?
  - Els que associen accés obert amb autoarxiu són una minoria
    - Més en física, matemàtiques, humanitats i art
    - Menys en medicina i ciències de la vida
  - Els favorables pensen en els beneficis:
    - Més lectors, més visibilitat
  - Són, essencialment, conservadors (No volen canvis)
  - Accés obert és una eina complementària, no reemplaça a les que ja existeixen

(Fry, Creaser, Whit,

Ciro Lluéca. Repositoris digitals. COBDC/BEG 2013. 2011)

### Cobertura estratègica (6)

- Beneficis d'un repositori:
  - Obrir i oferir els resultats de la institució o la comunitat al món
  - Maximitzar la visibilitat dels documents, proporcionant una major oportunitat d'impacte
  - Exhibir i vendre la institució a grups potencialment interessats: personal, estudiants, *partners*...
  - Donar suport al procés de recerca amb abundància de dades
  - Compilar i conservar la producció digital (i analògica, si existeix digitalització retrospectiva)
  - Proporcionar un escenari per al treball, de col·laboració i de projectes a gran escala

(Swan,

Ciro Lluéca. Repositoris digitals. COBDC/BEG 2013. 2011)

### Cobertura estratègica (7)

- El futur d'un repositori:
  - Identificar i captar continguts valuosos
  - Posar a disposició el contingut d'una manera sistemàtica i estandarditzada
  - Preservar els continguts
  - Ensenyar a trobar i a citar continguts
  - Educar sobre el *copyright* i la propietat intel·lectual
  - Ajudar en la publicació

(Hixson,  
2005)

Ciro Lluéca. Repositoris digitals. COBDC/BEG 2013.

### Repositori propi o cooperatiu (1)

- Repositori propi, si es tenen recursos, per incrementar i consolidar la visibilitat institucional
- Repositori cooperatiu per unir esforços i aconseguir ràpidament impacte dels autors i de la institució
- Aspectes a tenir en compte per prendre la decisió:
  - Política de la nostra institució
  - Usuaris
  - Característiques dels documents
- No són excloents!



Ciro Lluéca. Repositoris digitals. COBDC/BEG 2013.

## Repositori propi o cooperatiu (2)

Avantatges repositori propi	Avantatges repositori cooperatiu
Increment i consolidació de la visibilitat Institucional	Incorporació immediata en un repositori en funcionament amb els processos establerts
Personalització, autonomia, control, rapidesa i flexibilitat per experimentar o desenvolupar funcions ad-hoc	Agrupació dels continguts (més fàcil la recuperació de la informació per als usuaris i menor possibilitat de duplicar digitalitzacions)
Més fàcil convèncer als autors per a l'autoarxiu (més personalitzat)	Més visibilitat i difusió dels documents (més massa crítica, més usuaris, més consultes, millor posicionament en els cercadors, marca consolidada, etc.)
Eina per a la institució per analitzar i avaluar la seva recerca	Serveis de valor afegit que difícilment es poden tenir localment, per exemple, solucions costoses de preservació o migració o emulació de formats, últimes tendències tecnològiques, etc.
Millor posicionament en el Ranking Web of World repositories	Garantia de procediments establerts amb experiència, <i>know-how</i> i suma de coneixements
	Menys inversió econòmica (economia d'escala, racionalització i estalvi de costos)

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

- Identifica, en els documents de planificació estratègica de la teva pròpia unitat d'informació, els plans o estratègies en les que es podria incloure un projecte de repositori.
- Identifica còmplices en la teva pròpia institució i en el teu entorn: qui pot ajudar-te a finançar un nou repositori?

10' treball individual + 5'



## Cobertura legal (1)

- Les lleis de propietat intel·lectual vigents en la majoria dels països obliguen a obtenir permisos dels propietaris dels drets morals i/o materials per publicar en els repositoris, i en quins termes
- Més fàcil demanar perdó que demanar permís...!



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Cobertura legal (2)

- La propietat intel·lectual d'una obra literària, artística o científica correspon al seu autor pel sol fet de la seva creació
- Dos tipus de drets:
  - Personal
  - Patrimonial
- Serà necessari tenir l'autorització de l'autor o dels seus hereus per incloure les seves obres en el repositori

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Cobertura legal (3)

- Habitualment, els drets de les obres creades per treballadors d'una organització com a part de la seva feina, pertanyen a la organització:
  - Centres de recerca
  - Centres docents
  - Museus i centres culturals
  - Universitats
  - Biblioteques



Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

## Cobertura legal (4)

- A Espanya existeixen obres no subjectes a drets d'explotació:
  - Obres d'autor conegut mort > 80 anys
  - Fonogrames, audiovisuals i emissions radiofòniques > 50 anys
  - Fotografies > 25 anys
- Els drets morals no caduquen mai

(Ferrer et al., 2005, en base a la legislació

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

### Cobertura legal (5)

- En les obres orfes no es pot localitzar al titular dels drets:
  - Cerca diligent (criteris a unificar)
  - Bases de dades, catàlegs comercials
  - Llicències col·lectives per països (agències de drets)



“La tasca d’identificar i contactar amb els titulars dels drets representa el cost més alt en els projectes de preservació digital. La majoria ni responen a la sol·licitud”. (Laura Campbell, Library of Congress, 2006, citada per Keefer & Gallart 2007)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Cobertura legal (6)

- Els professionals disposem de bons i suficients arguments per convèncer als nostres futurs dipositants de la bona fe dels repositoris

(Vives, 2005)

Why is the Internet Archive collecting sites from the Internet? What makes the information useful?

Most societies place importance on preserving artifacts of their culture and heritage. Without such artifacts, civilization has no memory and no mechanism to learn from its successes and failures. Our culture now produces more and more artifacts in digital form. The Archive's mission is to help preserve those artifacts and create an Internet library for researchers, historians, and scholars. The Archive collaborates with institutions including the Library of Congress and the Smithsonian.

¿Por qué el Internet Archive colecciona los sitios web de Internet?

La mayoría de las sociedades dan importancia a las herramientas que logran la preservación de su cultura y su patrimonio. Sin ello, la civilización no tiene memoria y no existe un mecanismo para aprender de sus éxitos y fracasos. (...) La misión (...) es crear una biblioteca de Internet para investigadores, historiadores y eruditos



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Cobertura legal (7)

- En els repositoris institucionals és habitual que s'arxivin versions *pre o post-print* d'articles publicats en mitjans acadèmics
- Habitualment les revistes acadèmiques permeten aquesta pràctica
- L'autor haurà d'expressar per escrit el seu consentiment per al:
  - Dipòsit (tant si és autoarxiu com delegat)
  - Tranformació tècnica per a preservació
  - Comunicació pública a Internet



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Pràctica

- Visita **Dulcinea** (<http://www.accesoabierto.net/dulcinea/>) o **Sherpa/RoMEO** (<http://www.sherpa.ac.uk/romeo/>), són projectes per conèixer les polítiques editorials de les revistes en relació a l'autoarxiu en repositoris.
- Dulcinea** té un sistema de classificació segons la revista permeti pre i post (verd); post (blau); pre (verd) o cap opció (blanc). Cada revista es descriu segons la seva política.

verde azul amarillo blanco

10' treball individual



### Cobertura legal (8)

- En els documents nascuts digitals que s'inclouen en els repositoris d'accés obert s'estan usant ampliament les llicències *Creative Commons*
- Amb aquestes, l'autor determina a priori els usos que es poden fer de la seva obra
- A partir de 4 condicions s'elaboren 6 llicències, des de la més permissiva a la més restrictiva

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Cobertura legal (9)



CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

CC BY NC ND

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Cobertura legal (10)

- És recomanable que cada objecte dipositat en el repositori inclogui de manera clara les possibilitats d'ús del document
- Existeix la possibilitat de l'embargament (temporal, per IP, etc.)

<http://www.zaragoza.es/ciudad>

Condiciones de uso:  
Las imágenes contenidas en esta página web están libremente disponibles para su visualización.  
Para cualquier uso de las imágenes que implique reproducción con fines de publicación y difusión se deberá contar con la autorización expresa del Ayuntamiento de Zaragoza que pueden solicitar en: [archivo@zaragoza.es](mailto:archivo@zaragoza.es).  
El Ayuntamiento de Zaragoza no se responsabiliza del uso que puedan hacerse de las reproducciones en contra de los derechos de autor protegidos por la ley, estando sujetos a lo estipulado por la Ley de Propiedad Intelectual, así como a las normas del ordenamiento jurídico vigente en materia de acceso a los registros y archivos públicos, protección del derecho al honor, a la intimidad personal y familiar y a la propia imagen, así como a la ley de protección de datos de carácter personal, y demás legislación vigente.  
El acceso a las imágenes de esta web implica la aceptación de estas condiciones de uso.



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## 2.2 Maquinari i programari

### Maquinari (1)

- L'elecció entre un sistema local o extern depèn de diversos factors, bàsicament econòmics
  - Si la institució no té un equip informàtic estable i personal tècnic, serà més ràpid i econòmic ubicar el repositori en un sistema extern
  - Però, a llarg plaç, serà més car mantenir una estructura externa

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Maquinari (2)

	Servidor local	Servidor extern
Avantatges	<p>Velocitat en el tractament dels documents digitals</p> <p>Possibilitat d'aprofitar els recursos (maquinari, programari, personal) existents</p> <p>Treball més proper al personal tècnic responsable dels sistemes</p> <p>Més control i flexibilitat, més oportunitats per experimentar o desenvolupar funcions ad-hoc</p>	<p>Sense responsabilitat sobre els equips informàtics i el seu manteniment i garantia de més disponibilitat i control d'errors</p> <p>El sistema està operatiu des de l'inici, amb el conseqüent estalvi d'energia i temps en configuració, instal·lació, etc.</p> <p>Possibilitat de sumar-se a un consorci o acord cooperatiu existent per abaratir costos</p> <p>Si es treballa cooperativament, possibilitat de gestió més eficaç ja que es treballa en una empresa "neutral"</p>

(Basat en Clobridge, 2011)

### Maquinari (3)

	Servidor local	Servidor extern
Desavantatges	<p>Requereix personal d'administració de sistemes</p> <p>Requereix compra d'infraestructura tecnològica (maquinari i programari), manteniment i actualitzacions</p> <p>Requereix responsabilitzar-se de processos de back-up</p>	<p>Costos anuals</p> <p>Dificultats en cas de voler canviar de proveïdor del servidor</p> <p>Falta de control</p> <p>Ús de banda ample a Internet</p> <p>Possibles problemes derivats de la propietat intel·lectual per emmagatzemar la informació en servidors externs</p> <p>Possibles problemes relacionats amb la protecció de dades personals</p>

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Maquinari (4)

- Un exemple de sistema bàsic d'arquitectura tècnica local per un repositori:
  - Un node de la sèrie Proliant DL360, amb dos processadors de 6 nuclis més 24 GB RAM (ampliables a 128 GB) i un disc dur intern de 300 GB (ampliables a 2 TB)
  - L'exemple, d'HP, té un preu aprox. de 7.000 €. Altres versions de Sun o Dell consultar



HP Proliant DL360 G3, HP Proliant DL360 G4, HP Proliant DL360 G5, HP Proliant DL360 G6, HP Proliant DL360 G7

<http://h10010.www1.hp.com/wwpc/us/en/sm/WF04a/15351-15351-3328412-241644-241475.html>

### Maquinari (5)

- Altres factors que determinaran el grau de complexitat de l'arquitectura són:
  - Volum i tipus de documents a dipositar
  - Consultes d'usuaris potencials
- A més, del servidor, serà necessari disposar d'estacions de treball (escàners, càmeres...)
- Finalment, equips d'alimentació ininterrompuda i dispositiu de *back-up*

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### MDC, Memòria Digital de Catalunya

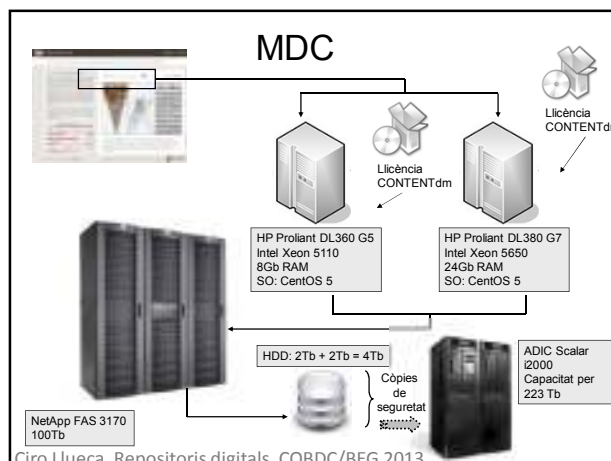


<http://www.cbuc.cat>

### MDC

- La Memòria Digital de Catalunya (2006- ) és un repositori cooperatiu que dona accés a col·leccions digitalitzades d'institucions científiques, culturals i/o erudites de Catalunya. Està impulsat per les universitats catalanes i la Biblioteca de Catalunya
- Funciona amb CONTENTdm, 2 milions de documents en 68 col·leccions
- Gran volum de documents i pes i molts usuaris potencials → Usa un únic node de 2TB. És la major col·lecció del món en CONTENTdm

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Pràctica

- Una de les col·leccions de la MDC és la dedicada al naturalista Dr. Jordi Sabater Pi (1922-1999), famós pel descobriment per a la ciència europea del Floquet de Neu, en 1966.
- Visita la seva col·lecció:
  - (<http://mdc.cbuc.cat/cdm4/browse.php?CISOROOT=/sabaterpi>):
  - Comprova les opcions de visualització i edició dels registres.

5' treball



### Programari (1)

- La tendència mundial és implementar els repositoris amb programari lliure
- Motius:
  - Adaptabilitat ràpida a cada cas
  - Costos que no inclouen llicències
  - Escrutini públic dels codis
  - Comunitats internacionals importants (continuitat del programari i millores)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Programari (2)

- L'elecció del programari es produeix en base al tipus de document a dipositar
- No existeix un únic programari amb solucions adients per totes les tipologies de documents
  - Les imatges, generalment més pesades, han de poder ser editades, sensibles a zoom, amb metadades visibles
  - Els manuscrits han d'incloure a més la transcripció del text
  - Les revistes digitals, els treballs acadèmics i documents textuais han de permetre la cerca en el text lliure

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Programari (3)

- Existeixen uns 70 programaris en ús
- Els més comuns són DSpace (27%) i EPrints (16%)
  - Són fàcils d'instal·lar i mantenen i gestionen tot tipus de formats
  - Són de codi obert (OpenDoar.org, 2011)
  - A Espanya el 68% dels repositoris usen DSpace (Abadal et al., 2010)
- Fedora és el més complex
  - Requereix instal·lació i desenvolupament tècnic expert
  - També és de codi obert
- Greenstone és molt usat per institucions amb pocs recursos
  - De codi obert
  - És senzill d'instal·lar i mantenir
- CONTENTdm, d' OCLC, ideal per col·leccions de mida limitada

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Programari (4)

- És recomanable que els sistemes de repositori es construeixin sobre sistemes de bases de dades, com Oracle, Microsoft SQL, o MySQL, per garantir la flexibilitat a llarg plaç
- Existeixen poques comparatives públiques i guions per l'elecció del sistema:
  - Singh et al. A comparative analysis of institutional repository software. *International Conference on Open Repositories* 2010.  
[http://biecoll.ub.uni-bielefeld.de/volltexte/2011/5076/pdf/abs\\_singh\\_comparative.pdf](http://biecoll.ub.uni-bielefeld.de/volltexte/2011/5076/pdf/abs_singh_comparative.pdf)
  - Open Society Institute. *A guide to institutional repository software* 2004  
[http://www.soros.org/openaccess/pdf/OSI\\_Guide\\_to\\_IR\\_Software\\_v3.pdf](http://www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## OJS

- OJS (Open Journal Systems), desenvolupat pel Public Knowledge Project (<http://pkp.sfu.ca>) és un programari lliure per la gestió i publicació de publicacions en sèrie



<http://pkp.sfu.ca/>

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Revistas Científicas Complutenses



<http://revistas.ucm>

## Revistas Científicas Complutenses

- La Universidad Complutense de Madrid és una universitat pública amb seu a Madrid. Es considera la successora de la fundada al 1499 a Alcalá de Henares. És la universitat espanyola amb un major nombre d'estudiants presencials (85.000 al curs 2010-11). Ofereix 180 titulacions
- La Biblioteca Complutense té un pressupost anual de 1,3 M€ (2010). Al 2008 era de 2,4 M€
- El portal de Revistas Científicas Complutenses és un repositori institucional que inclou els articles científics de les publicacions realitzades per la pròpia universitat
- OJS, 9.900 registres



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



## Pràctica

- En el portal de *Revistas Científicas Complutenses* (<http://revistas.ucm.es>):
  - Localitza a la "Búsqueda avanzada" articles sobre campanyes electorals. Revisa les possibilitats de visualització de resultat.
  - Accedeix a la fitxa de la revista "Anaquel de Estudios Arabes". Quina part de la col·lecció és subjecta a embargament? Si és possible, accedeix al text complet d'un dels articles.

5' treball



## CONTENTdm

- CONTENTdm és un programari propietari que comercialitza DOC6 (OCLC).
- S'usa majoritàriament en repositoris d'imatges, però administra tot tipus de formats
- A Espanya el cost per llicència és d'uns 6.000 €


<http://www.contentdm.org>

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Fundació Joan Miró


<http://www.fundaciomiro.org>

## Fundació Joan Miró

- La Fundació Joan Miró va obrir al públic l'any 1975. De titularitat privada amb seu a Barcelona, té com objectiu difondre i protegir l'obra de l'artista Joan Miró (1893-1983)
- El seu museu compren més de 14.000 peces de l'artista, així com d'altres contemporanis (Duchamp, Calder, Chillida)
- La col·lecció de la fundació és accessible mitjançant una senzilla classificació per tècniques i època cronològica, així com per cerca avançada en el text lliure
- El repositori facilita la divulgació de la Fundació, i també en les sales del museu, mitjançant pantalles tàctils
- CONTENTdm, 1.000 registres



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

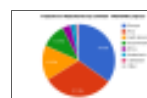
- En el fons documental de la Fundació Joan Miró (<http://80.25.201.238:8081/cdm-fons-fjm/search.php?CISOROOT=/fons-fjm>):
  - Localitza a la "Búsqueda avanzada" documents que continguin la combinació dels conceptes "dona" i "ocell". Identifica les possibilitats de gestió dels resultats.
  - Existeix algun dibuix de Mallorca?

5' treball



## DSpace

- DSpace és un programari de codi lliure massivament usat a Espanya i a la resta del món
- Creat inicialment pel MIT (Massachusetts Institute of Technology) i Hewlett Packard, actualment és la comunitat internacional la que manté actiu el sistema
- Destinat principalment a documentació textual (articles, tesis, comunicacions...).


<http://www.dspace.org>

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

DSpace

## Memoria Digital Vasca



<http://www.memoriadigitalvasca.es/>

## Memoria Digital Vasca

- La Fundación Sancho El Sabio va ser creada l'any 1955, té des de 2009 l'objectiu de difondre la cultura vasca en *open access*
- El repositori Memoria Digital Vasca (=Euskal Memoria Digitala) conté manuscrits, mapes, material gràfic, monografies i revistes, descrits i classificats en base a un sistema propi de 1.400 epígrafs
- DSpace, 3.000 registres



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

- En el repositori Memoria Digital Vasca (<http://www.memoriadigitalvasca.es/>):
  - Localitza el document “Don Francisco Zazo de Ulloa...” i accedeix a les dades descriptives, al document complet i a les imatges de mostra.

5' treball

individual

## RUA, Repositori Institucional de la Universitat d'Alacant



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

<http://rua.ua.es/>

## RUA

- La Universitat d'Alacant és una universitat pública creada l'any 1979
- 28.000 estudiants al 2009 / 50 titulacions
- La Biblioteca de la UA gestiona el repositori RUA, amb l'objectiu de recopilar, gestionar i preservar la documentació digital generada pels membres de la universitat i per promoure l'accés obert als resultats de la seva recerca
- Conté tot tipus de formats (articles, *preprints*, *working papers*, informes tècnics, comunicacions a congressos, llibres, tesis, imatges, etc.)
- DSpace, 18.000 registres



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

- En el RUA (<http://rua.ua.es/>):
  - A “Estadísticas”, detecta quins documents són els més visitats del repositori.
  - Quin tipus de document ha dipositat Antonia Montes Fernández?
  - Localitza el video en el que es demostra el principi d'Arquímedes.

5' treball

individual

## Fedora

- Fedora (Flexible Extensible Digital Object Repository Architecture) és un programari lliure creat l'any 1997 per diverses institucions americanes
- El sistema permet gestionar col·leccions de documents complexes (vídeo, audio, bases de dades), a més de convencionals (articles, etc.)
- Al 2009 Fedora i DSpace van anunciar la seva fusió, a DuraSpace



<http://fedora->

[FedoraCommons](http://fedora-commons.org/)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Repositori de la UNED



<http://e-spacio.uned.es>

## Universidad Nacional de Educación a Distancia (UNED)

- La UNED (Universidad Nacional de Educación a Distancia) és una universitat pública d'àmbit estatal. Creada el 1973, ofereix estudis a distància
- Al 2010 = 200.000 estudiants i 227 M€ de pressupost
- El repositori de la UNED conté els articles de les revistes digitals de la universitat, manuals escolars i recursos educatius, actes de congressos propis i els fons antics digitalitzats
- Fedora, 12.000 registres



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

- Al Repositori de la UNED (<http://e-spacio.uned.es>):
  - Accedeix a documents dels "Fondos digitalizados" per consultar el tipus de fitxa descriptiva, les opcions de visualització i les opcions de metadades.

5' treball

individual

## EPrints

- EPrints és un programari lliure dissenyat al 2000 per contenir dipòsits digitals de recerca universitària
- Darrera de DSpace, és un programari àmpliament usat que admet tot tipus de formats



<http://www.eprint>

[eprints](http://www.eprint.org/)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## RIBEI



<http://ribei.org/>



### Programari (5)

- Si el repositori conté únicament imatges, i és per ús personal, domèstic, o d'un centre amb mínims recursos, pot ser una bona idea usar alguns dels sistemes gratuïts, com Picasa, Flickr, o Omeka



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Pràctica

- Visita a Flickr el perfil de la *Biblioteca de la Facultad de Derecho y Ciencias del Trabajo Universidad de Sevilla*: (<http://www.flickr.com/people/fdctsevilla/>)
- Visita la Biblioteca digital de Roquetes: (<http://bibliotecaroquetes.cat/>)

5' treball individual

### Programari (6)

- Són nombroses les institucions que han creat programari a mida:
  - Per desconeixement del mercat
  - Per voluntat de comptar amb una arquitectura específica
  - Per usar eines senzilles que poden servir per un repositori senzill



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Programari (7)



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Pràctica

- Visita un dels tres recursos mencionats i valora'n llur funcionament.
  - ADACYL,  
<http://www.adacyl.org/>
  - CTPedia,  
<http://www.ctpedia.es/>
  - Memoria de Madrid,  
<http://www.memoriademadrid.es>

5' treball individual

### Pràctiques amb Dspace i CONTENTdm

### 3.1. Planificació dels processos de treball

#### Planificació (1)

- Habitualment el personal en plantilla és el que es fa càrrec de la creació i manteniment del repositori
- És convenient detallar les tasques a dur a terme en el procés:
  - Coordinació
  - Metadades i càrrega d'objectes
  - Suport informàtic
  - Legal, Publicacions, Gestió acadèmica...



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

#### Planificació (2)

##### Planificació de tasques

**Pla operatiu:** implementar un repositori digital al servei de la institució o participar en un repositori ja existent

**Actuacions prèvies:** identificació i digitalització de la col·lecció segons les polítiques i els paràmetres establerts, consideració dels aspectes legals, selecció del sistema de metadades, implementació del maquinari i programari (si hem decidit crear un repositori propi), selecció dels serveis a oferir a l'usuari i assignació dels recursos humans

**Actuació 1:** ordenació de la col·lecció digital en un sistema de fàcil accés (tècnic)

**Actuació 2:** configuració dels detalls de la col·lecció en el repositori i dels perfils d'usuari (administrador/a)

**Actuació 3:** càrrega d'objectes digitals i assignació de metadades (tècnic)

**Actuació 4:** validació de la càrrega i publicació en obert (administrador/a)

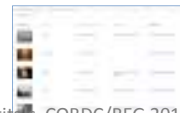
**Actuació 5:** promoció del repositori (coordinador/a del repositori)

**Actuació 6:** avaluació del repositori (coordinador/a del repositori)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

#### Planificació (3)

- La càrrega de documents digitals i la descripció dels mateixos depèn del programari del repositori
  - DSpace o EPrints treballen directament en un servidor extern
  - Amb CONTENTdm es treballa en local, en l'estació de treball de l'operador



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

#### Planificació (4)

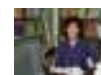
- Un administrador és el responsable de:
  - Definir les col·leccions (per formats, per temàtiques, per procedències)
  - Crear els perfils dels operadors ("submitters" a DSpace)
  - Preparar els permisos dels usuaris (accés lliure, restringit)
  - Dissenyar la interfície de consulta (imatge corporativa)



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

#### Planificació (5)

- Un operador és el responsable de:
  - Organitzar els fitxers (per pàgines, per documents)
  - Carregar els arxius
  - Assignar les metadades (administratives, descriptives, de preservació) o importar-les des d'altres catàlegs
  - Carregar l'objecte al servidor, supervisió oportuna



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Hispana



<http://hispana.mcu.es>

## Hispana

- Hispana (Ministerio de Cultura de España, 2010) reuneix les col·leccions digitals d'arxius, biblioteques i museus conformes a la iniciativa d'arxiu oberts que promou la Unió Europea. És, de facto, la *Europeana* espanyola
- És un agregador de continguts de 174 repositoris espanyols d'universitats, biblioteques digitals, museus, etc. Dóna accés a 3,4 milions de registres
- També manté un directori de 568 projectes de digitalització

Ciro Lluëca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

- Accedeix a **Hispana** (<http://hispana.mcu.es>):
- Busca per "Margarita Xirgu" per comprovar les opcions de refinament de resultats i la visualització de metadades.
- La plaça de braus de "Las Arenas" de Barcelona ha estat convertida en un centre comercial. Busca fotos antigues de la plaça.
- Localitza un enregistrament que commemori la mort del torero Manolete (1917-1947).
- Accedeix a cartes manuscrites de la família Bonaparte.

10' treball individual

## 3.2. Publicació i promoció

### Publicació (1)

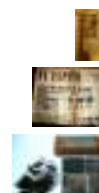
- En paral·lel a comptar amb cobertura estratègica i legal, i a definir les necessitats de maquinari i programari, és necessari:
  - Definir les col·leccions que formaran part del repositori (pensant en clau d'impacte)
  - Identificar els recursos necessaris per a la implementació
  - Determinar el sistema de metadades

Ciro Lluëca. Repositoris digitals. COBDC/BEG 2013.



### Publicació (2)

- En el cas de col·leccions analògiques serà necessari un projecte previ de digitalització, que prioritzi:
  - Materials de major interès per la institució
  - Fragilitat
  - Singularitat
  - Interès per a la investigació (transcripcions, etc.)
  - Freqüència d'ús i consulta
  - Domini públic
  - Col·leccions completes
  - Requeriments especials de visualització
- L'ha digitalitzat alguna altra institució?



Ciro Lluëca. Repositoris digitals. COBDC/BEG 2013.



### Publicació (3)

- Un projecte de digitalització no és econòmic:
  - Tesi doctoral de 450 pàgines, 88€
  - Revista enquadernada en color de 50 pàgines, 26€
  - Revista enquadernada en blanc i negre de 50 pàgines, 15€
  - Col·lecció de 1.000 documents amb OCR, lliurament amb doble còpia (difusió + preservació), 650€
  - Col·lecció de 1.000 fotografies antigues, digitalitzades, descrites i carregades al repositori, 1.700€
- Comptem amb patrocinis o subvencions?

Ciro Ilueca. Repositoris digitals. COBDC/BEG 2013.



### BV Miguel de Cervantes



<http://www.cervantesvirtual.com/>

### BV Miguel de Cervantes

- La Universitat d'Alacant es va plantejar al 1998 aplicar els seus coneixements tecnològics per crear amb un equip de filòlegs i informàtics una "biblioteca digital" de literatura en espanyol
- Aquesta va ser inaugurada al juny de 1999 amb 3.000 llibres digitalitzats
- Amb l'èxit obtingut, es va iniciar una fórmula de col·laboració pública-privada que ha resultat ser modèlica. Avui en dia és una fundació que ofereix en accés obert 130.000 documents de diverses biblioteques
- La col·lecció de difusió, dipositada en servidors d'Amazon, es complementa amb les còpies de seguretat per preservació, en la Universitat d'Alacant

**Manuel Bravo** és el seu director:  
"Queremos un portal de literatura en español accesible para todo el mundo."



### Pràctica

- Visita la Biblioteca Virtual Miguel de Cervantes (<http://www.cervantesvirtual.com/>):
  - Busca obres clàssiques (La Celestina, p.e.).
  - Busca obres d'un/a autor/a contemporani del teu interès.
  - Revisa el sistema de funcionament i accés al document en les cerques anteriors.
  - Accedeix a la secció de "Páginas de autor".

10' treball individual

### Publicació (4)

- Un repositori no és tal si no compta amb un número significatiu de documents i un creixement permanent
- En el disseny de la estratègia d'implementació del repositori:
  - Prova pilot
  - La producció de les primeres col·leccions
  - L'impacte del repositori en la comunitat

Ciro Ilueca. Repositoris digitals. COBDC/BEG 2013.



### Publicació (5)

- Fer sostenibles els repositoris és l'objectiu:
  - Que s'alimentin de manera autosuficient
    - Les polítiques afecten al comportament: el 62% dels autoarxius són per obligació; només el 20% són voluntaris (Alma Swan, 2011)
  - Que comptin amb financiació permanent
  - Que siguin usats i reconeguts
    - D'ús i re-ús lliure; d'accés immediat; connectats i connectables; amigables, fàcils d'usar (Alma Swan, 2011)

Ciro Ilueca. Repositoris digitals. COBDC/BEG 2013.





## RiuNet



<http://riunet.upv.es/>

## RiuNet

- La Universitat Politècnica de València (UPV) és una universitat pública creada al 1968, amb seu a València. Té 2.855 professors i 36.000 estudiants (2010). Els seu pressupost anual és de 373 M€ (2011)
- El seu repositori, RiuNet, té l'objectiu d'arxivar la producció personal o institucional de la seva comunitat
- És exemplar pel volum del seu contingut educatiu: el 59% són d'aprenentatge i/o docència (Bueno, Hernández, 2011)
- DSpace, 7.771 documents



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

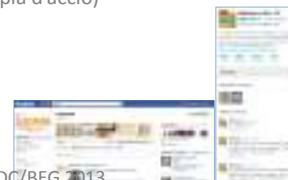
- Visita RiuNet (<http://riunet.upv.es/>):
  - Begoña Ivars ha publicat diversos objectes d'aprenentatge. Visita aquesta col·lecció i comprova el seu sistema de funcionament.
  - Luis Ángel Alonso ha dipositat diverses presentacions que empra en les seves classes. Visita els seus passejos per Mart.
  - Ulises Plancarte ha dipositat una interessant tesi. ¿Quina és?

5' treball

individual

## Promoció (1)

- Les estratègies de comunicació han de considerar als dipositants potencials, així com als usuaris
- Establir aliances és imprescindible: directius, departaments paral·lels, unitats d'informació, altres repositoris, etc.
- Ús de la web social (amb un pla d'acció)



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Promoció (2)

- Eines de promoció:
  - Contingut unit orientat a usuaris finals, amb materials de formació, instruccions i polítiques
  - Presentació clara amb els beneficis, per als investigadors i professors *senior*
  - Tallers a mida per fomentar l'ús del repositori
  - Arguments clars amb els beneficis per als usuaris finals (diferents rols / FAQs)
  - Enllaços incrustats al repositori des de la intranet i extranet de la universitat
  - Difusió de la informació (alertes, RSS, novetats, xarxes socials) en els canals de comunicació de la institució

(Holland, Denning,

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013 (2011)

## Promoció (3)

- Motius per NO usar el repositori:
  - Varietat de repositoris, confusió
  - Falta de temps per aprendre aplicacions
  - Confusió sobre els drets d'autor
  - Por al plagi
  - Por a revelar l'estat de la investigació
  - Percepció psicològica de baixa qualitat dels materials dipositats
  - Dubte sobre si el material dipositat en el repositori es considera publicat

(Davis, Connolly,

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013 (2007)

### Promoció (4)

- Incentius per al reclutament de continguts:
  - Reconeixement públic dels autors/usuaris resultant la seva tasca (lístes d'autors que més dipositen, events d'innovació docent, butlletins i memòries, etc.)
  - Reconeixement formal mitjançant registre i expedició de certificats dels repositoris
  - Reconeixement com mèrits d'investigació
  - Reconeixement a la qualitat i al valor dels seus materials (contingut, innovació, etc.)
  - Incentius econòmics als docents més actius i amb recursos educatius més populars i usats
  - Incentius econòmics per mèrits en la participació en el repositori

(Bueno, Hernández,

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013)

### Promoció (5)

- Incentius per al reclutament de continguts:
  - Serveis per minimitzar l'esforç dels autors (captar continguts des dels gestors de recerca, digitalitzar versions impreses, generar metadades, etc.)
  - Informar als autors (estadístiques de visualització, descàrrega, impacte)
  - Assessorar sobre propietat intel·lectual

(Keefer, 2008; basado en Proudman 2008, Royster 2007)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Promoció (6)

- La visibilitat professional del repositori:
  - Directoris: OAI Data providers, OpenDOAR, Buscarepositorios
  - Agregadors: DRIVER, OpenAire, DART, OAISter, Europeana, Hispana, Recolecta, Eureka, etc.
  - Ranking mundial de repositoris

For more information on OAI Data Providers, visit the [www.openarchives.org/Registry/BrowseSites](http://www.openarchives.org/Registry/BrowseSites) website. Currently there are 108 OAI Data Providers. The table may be sorted by the number of records, the number of records in the last 24 hours, the number of records in the last 7 days, the number of records in the last 30 days, the number of records in the last 90 days, the number of records in the last 180 days, the number of records in the last 360 days, the number of records in the last 720 days, the number of records in the last 1440 days, the number of records in the last 2880 days, the number of records in the last 5760 days, the number of records in the last 11520 days, the number of records in the last 23040 days, the number of records in the last 46080 days, the number of records in the last 92160 days, the number of records in the last 184320 days, the number of records in the last 368640 days, the number of records in the last 737280 days, the number of records in the last 1474560 days, the number of records in the last 2949120 days, the number of records in the last 5898240 days, the number of records in the last 11796480 days, the number of records in the last 23592960 days, the number of records in the last 47185920 days, the number of records in the last 94371840 days, the number of records in the last 188743680 days, the number of records in the last 377487360 days, the number of records in the last 754974720 days, the number of records in the last 1509949440 days, the number of records in the last 3019898880 days, the number of records in the last 6039797760 days, the number of records in the last 12079595520 days, the number of records in the last 24159191040 days, the number of records in the last 48318382080 days, the number of records in the last 96636764160 days, the number of records in the last 193273528320 days, the number of records in the last 386547056640 days, the number of records in the last 773094113280 days, the number of records in the last 1546188226560 days, the number of records in the last 3092376453120 days, the number of records in the last 6184752906240 days, the number of records in the last 12369505812480 days, the number of records in the last 24739011624960 days, the number of records in the last 49478023249920 days, the number of records in the last 98956046499840 days, the number of records in the last 197912092999680 days, the number of records in the last 395824185999360 days, the number of records in the last 791648371998720 days, the number of records in the last 1583296743997440 days, the number of records in the last 3166593487994880 days, the number of records in the last 6333186975989760 days, the number of records in the last 12666373951979520 days, the number of records in the last 25332747903959040 days, the number of records in the last 50665495807918080 days, the number of records in the last 101330991615836160 days, the number of records in the last 202661983231672320 days, the number of records in the last 405323966463344640 days, the number of records in the last 810647932926689280 days, the number of records in the last 1621295865853378560 days, the number of records in the last 3242591731706757120 days, the number of records in the last 6485183463413514240 days, the number of records in the last 12970366926827028480 days, the number of records in the last 25940733853654056960 days, the number of records in the last 51881467707308113920 days, the number of records in the last 103762935414616227840 days, the number of records in the last 207525870829232455680 days, the number of records in the last 415051741658464911360 days, the number of records in the last 830103483316929822720 days, the number of records in the last 1660206966633859645440 days, the number of records in the last 3320413933267719290880 days, the number of records in the last 6640827866535438581760 days, the number of records in the last 13281655733070877163520 days, the number of records in the last 26563311466141754327040 days, the number of records in the last 53126622932283508654080 days, the number of records in the last 106253245864567017308160 days, the number of records in the last 212506491729134034616320 days, the number of records in the last 425012983458268069232640 days, the number of records in the last 850025966916536138465280 days, the number of records in the last 1700051933833072276930560 days, the number of records in the last 3400103867666144553861120 days, the number of records in the last 6800207735332289107722240 days, the number of records in the last 13600415470664578215444480 days, the number of records in the last 27200830941329156430888960 days, the number of records in the last 54401661882658312861777920 days, the number of records in the last 108803323765316625723555840 days, the number of records in the last 217606647530633251447111680 days, the number of records in the last 435213295061266502894223360 days, the number of records in the last 870426590122533005788446720 days, the number of records in the last 1740853180245066011576893440 days, the number of records in the last 3481706360490132023153786880 days, the number of records in the last 6963412720980264046307573760 days, the number of records in the last 13926825441960528092615147520 days, the number of records in the last 27853650883921056185230295040 days, the number of records in the last 55707301767842112370460590080 days, the number of records in the last 111414603535684224740921180160 days, the number of records in the last 222829207071368449481842360320 days, the number of records in the last 445658414142736898963684720640 days, the number of records in the last 891316828285473797927369441280 days, the number of records in the last 1782633656570947595854738882560 days, the number of records in the last 3565267313141895191709477765120 days, the number of records in the last 7130534626283790383418955530240 days, the number of records in the last 14261069252567580766837911060480 days, the number of records in the last 28522138505135161533675822120960 days, the number of records in the last 57044277010270323067351644241920 days, the number of records in the last 114088554020540646134703288483840 days, the number of records in the last 228177108041081292269406576967680 days, the number of records in the last 456354216082162584538813153935360 days, the number of records in the last 912708432164325169077626307870720 days, the number of records in the last 1825416864328650338155252615741440 days, the number of records in the last 3650833728657300676310505231482880 days, the number of records in the last 7301667457314601352621010462965760 days, the number of records in the last 14603334914629202705242020925931520 days, the number of records in the last 29206669829258405410484041851863040 days, the number of records in the last 58413339658516810820968083703726080 days, the number of records in the last 116826679317033621641936167407452160 days, the number of records in the last 233653358634067243283872334814904320 days, the number of records in the last 467306717268134486567744669629808640 days, the number of records in the last 934613434536268973135489339259617280 days, the number of records in the last 1869226869072537946270978678519234560 days, the number of records in the last 3738453738145075892541957357038469120 days, the number of records in the last 7476907476290151785083914714076938240 days, the number of records in the last 14953814952580303570167829428153876480 days, the number of records in the last 29907629905160607140335658856307752960 days, the number of records in the last 59815259810321214280671317712615505920 days, the number of records in the last 119630519620642428561342635425231011840 days, the number of records in the last 239261039241284857122685270850462023680 days, the number of records in the last 478522078482569714245370541700924047360 days, the number of records in the last 957044156965139428490741083401848094720 days, the number of records in the last 1914088313930278856981482166803696189440 days, the number of records in the last 3828176627860557713962964333607392378880 days, the number of records in the last 7656353255721115427925928667214784757760 days, the number of records in the last 15312706511442230855851857334429569515520 days, the number of records in the last 30625413022884461711703714668859139031040 days, the number of records in the last 61250826045768923423407429337718278062080 days, the number of records in the last 122501652091537846846814858675436556124160 days, the number of records in the last 245003304183075693693629717350873112248320 days, the number of records in the last 490006608366151387387259434701746224496640 days, the number of records in the last 980013216732302774774518869403492448993280 days, the number of records in the last 1960026433464605549549037738806844897986560 days, the number of records in the last 3920052866929211099098075477613689795973120 days, the number of records in the last 7840105733858422198196150955227379591946240 days, the number of records in the last 15680211467716844396392301910454759183892480 days, the number of records in the last 31360422935433688792784603820909518367784960 days, the number of records in the last 62720845870867377585569207641819036735569920 days, the number of records in the last 125441691741734755171138415283638073471139840 days, the number of records in the last 250883383483469510342276830567276146942279680 days, the number of records in the last 501766766966939020684553661134552293884559360 days, the number of records in the last 1003533533933878041369107322269104587769118720 days, the number of records in the last 2007067067867756082738214644538209175538237440 days, the number of records in the last 4014134135735512165476429289076418351076474880 days, the number of records in the last 8028268271471024330952858578152836702152949760 days, the number of records in the last 16056536542942048661905717156305673404305899520 days, the number of records in the last 32113073085884097323811434312611346808611799040 days, the number of records in the last 64226146171768194647622868625222693617223598080 days, the number of records in the last 128452292343536389295245737250445387234447196160 days, the number of records in the last 256904584687072778590491474500890774468894392320 days, the number of records in the last 513809169374145557180982949001781548937788784640 days, the number of records in the last 1027618338748291114361965898003563097875577569280 days, the number of records in the last 2055236677496582228723931796007126195751155138560 days, the number of records in the last 4110473354993164457447863592014252391502310277120 days, the number of records in the last 8220946709986328914895727184028504783004620554240 days, the number of records in the last 16441893419972657829791454368057009566009241108480 days, the number of records in the last 32883786839945315659582908736114019132018482216960 days, the number of records in the last 65767573679890631319165817472228038264036964433920 days, the number of records in the last 131535147359781262638331634944456076528073928867840 days, the number of records in the last 263070294719562525276663269888912153056147857735680 days, the number of records in the last 526140589439125050553326539777824306112295715471360 days, the number of records in the last 1052281178878250101106653079555648612224591429542720 days, the number of records in the last 2104562357756500202213306159111297224449182859085440 days, the number of records in the last 4209124715513000404426612318222594448898365718170880 days, the number of records in the last 8418249431026000808853224636445188897796731436341760 days, the number of records in the last 16836498862052001617706449272890377795593462872683520 days, the number of records in the last 33672997724104003235412898545780755591186925745367040 days, the number of records in the last 67345995448208006470825797091561511182373851490734080 days, the number of records in the last 134691990896416012941651594183123022364747702981468160 days, the number of records in the last 269383981792832025883303188366246044729495405962936320 days, the number of records in the last 538767963585664051766606376732492089458990811925872640 days, the number of records in the last 1077535927171328103533212753464984178917981623851745280 days, the number of records in the last 2155071854342656207066425506929968357835963247703490560 days, the number of records in the last 4310143708685312414132851013859936715671926495406981120 days, the number of records in the last 8620287417370624828265702027719873431343852990813962240 days, the number of records in the last 17240574834741249656531404055439746862687705981627924480 days, the number of records in the last 34481149669482499313062808110879493725375411963255848960 days, the number of records in the last 68962299338964998626125616221758987450750823926511697920 days, the number of records in the last 137924598677929997252251232443517974901501647853023395840 days, the number of records in the last 275849197355859994504502464887035949803003295706046791680 days, the number of records in the last 551698394711719989009004929774071896066006591412093583360 days, the number of records in the last 1103396789423439978018009859548143792132013182824187166720 days, the number of records in the last 2206793578846879956036019719096287584264026365648374333440 days, the number of records in the last 4413587157693759912072039438192575168528052731296748666880 days, the number of records in the last 8827174315387519824144078876385150337056105462593497333760 days, the number of records in the last 17654348630775039648288157752770300674112210925186994667520 days, the number of records in the last 35308697261550079296576315505540601348224421850373989335040 days, the number of records in the last 70617394523100158593152631011081202696448843700747978670080 days, the number of records in the last 141234789046200317186305262022162405392897687401495957340160 days, the number of records in the last 282469578092400634372610524044324810785795374802991914680320 days, the number of records in the last 564939156184801268745221048088649621571590749605983829360640 days, the number of records in the last 1129878312369602537490442096177299243143181499211967658721280 days, the number of records in the last 2259756624739205074980884192354598486286362998423935317442560 days, the number of records in the last 4519513249478410149961768384709196972572725996847870634885120 days, the number of records in the last 9039026498956820299923536769418393945145451993695741269770240 days, the number of records in the last 18078052997913640599847073538836787890290903987391482539540480 days, the number of records in the last 36156105995827281199694147077673575780581807974782965079080960 days, the number of records in the last 72312211991654562399388294155347151561163615949565930158161920 days, the number of records in the last 144624439983309124798776588310694303122327231899131860316323840 days, the number of records in the last 289248879966618249597553176621388606244654463798263720632647680 days, the number of records in the last 578497759933236499195106353242777212489308927596527441265295360 days, the number of records in the last 1156995519866472998390212706485544424978617855193054882530590720 days, the number of records in the last 2313991039732945996780425412971088849957235710386109761061181440 days, the number of records in the last 4627982079465891993560850825942177699914471420772219522122362880 days, the number of records in the last 9255964158931783987121701651884355399828942841544439044244725760 days, the number of records in the last 18511928317863567974243403303768710799657885683088878088489451520 days, the number of records in the last 37023856635727135948486806607537421599315771366177756176978903040 days, the number of records in the last 74047713271454271896973613215074843198631542732355512353957806080 days, the number of records in the last 148095426542908543793947226430149686397263085464711024707915612160 days, the number of records in the last 296190853085817087587894452860299372794526170929422049415831224320 days, the number of records in the last 592381706171634175175788905720598745589052341858844098831662448640 days, the number of records in the last 1184763412343268350351577811441197491178104683717688197663324897280 days, the number of records in the last 2369526824686536700703155622882394982356209367435376395326649794560 days, the number of records in the last 4739053649373073401406311245764789964712418734870752790653299589120 days, the number of records in the last 9478107298746146802812622491529579929424837469741505581306599178240 days, the number of records in the last 1895621459749229360562524498305915985884967493948301116261199356640 days, the number of records in the last 3791242919498458721125048996611831971769934987896602232522398713280 days, the number of records in the last 7582485838996917442250097993223663943539869975793204465044797426560 days, the number of records in the last 15164971677993834884500195986447327887079739951586408930089594853120 days, the number of records in the last 30329943355987669769000391972894655774159479903172817860179189706240 days, the number of records in the last 60659886711975339538000783945789311548318959806345635720358379412480 days, the number of records in the last 121319773423950679076001567891578623096637919612691271440716758824960 days, the number of records in the last 242639546847901358152003135783157246193275839225382542881433517649920 days, the number of records in the last 485279093695802716304006271566314523386551678450765085762867035299840 days, the number of records in the last 970558187391605432608012543132629046773103356901530171525734070599680 days, the number of records in the last 1941116374783210865216025086265258093546206713803060343051468141199360 days, the number of records in the last 3882232749566421730432050172530516187092413427606120686102936282398720 days, the number of records in the last 7764465499132843460864100345061032374184826855212241372205872564797440 days, the number of records in the last 15528930998265686921728200690122064748369653710424482744411745129594880 days, the number of records in the last 31057861996531373843456401380244129496739307420848965488823490259189760 days, the number of records in the last 62115723993062747686912802760488258993478614841697930977646980518379520 days, the number of records in the last 124231447986125495373825605520976517986957229683395861955293961036759040 days, the number of records in the last 248462895972250990747651211041953035973914459366791723910587922073518080 days, the number of records in the last 496925791944501981495302422083906071947828918733583447821175844147036160 days, the number of records in the last 99385158388900396299060484416

## Ranking Web de Repositorios del Mundo

- El CSIC (Centro Superior de Investigaciones Científicas) és la major agència pública espanyola d'investigació científica (3a UE), amb 15.000 persones en plantilla (d'aquestes 3.000 són investigadors), genera el 20% de la producció científica espanyola
- Va signar la Declaració de Berlín al 2006
- El Cybermetrics Lab es dedica a l'anàlisi quantitatiu de Internet i els seus continguts, especialment d'aquells relacionats amb el procés de generació i comunicació acadèmica del coneixement científic
- Gestiona múltiples projectes, entre els quals el Ranking Web de Universidades del mundo i el Ranking Web de Repositorios del mundo

**Isidro Aguillo** és el seu responsable:

"El objetivo de un repositorio no es conseguir más visibilidad para la institución o la biblioteca, sino para los artículos que alberga."



## Ranking Web de Repositorios del Mundo

- Estableix una metodologia pròpia, construïda amb indicadors de presència a Internet i d'impacte o visibilitat hipertextual en els grans cercadors (Google, Yahoo, Bing):
  - Mida: número de pàgines recuperades
  - Visibilitat: número total d'enllaços externs rebuts (inlinks)
  - Fitxers rics: número de fitxers .pdf, .doc, .ppt, i ps
  - Scholar: càlcul del número total i d'articles recents (2006- )
- Inclou repositoris amb domini propi (repositori.cat; no [www.xxx.cat/repositori](http://www.xxx.cat/repositori)) i de fonamentalment treballs científics

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Ranking Web de Repositorios del Mundo

- Decàleg de bones pràctiques de posicionament:
  - Assignació URL (marca única)
  - Crear continguts (centrals + biblioteques + autors)
  - Convertir continguts (informes, literatura gris, etc.)
  - Interconnexió (present en la teva comunitat i xarxes)
  - Anglès (web i continguts)
  - Fitxers rics i multimèdia (pdf, ppt, wav)
  - Disseny amigable per als cercadors (estètica vs pràctica)
  - Popularitat i estadístiques de qualitat (qui ens enllaça?)
  - Arxivatge i persistència (preservar els continguts antics)
  - Estàndards (Dublin Core, per exemple)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

- Accedeix al **Ranking Mundial de Repositorios** (<http://repositories.webometrics.info>):
- Què tenen en comú els repositoris millor situats en el rànking?
- Quin és el repositori europeu millor situat? Visita i valora'n el seu funcionament.

10' treball



### 3.3. Avaluació: visibilitat, interoperabilitat i polítiques públiques

## Avaluació (1)

- Existeixen diverses guies per avaluar els repositoris
- Són eines a tenir en compte abans, durant i després de la implementació
- Alguns dels aspectes a avaluar: la visibilitat, accés a les polítiques del repositori, aspectes legals, interoperabilitat, etc.



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Avaluació (2)

- Deutsche Initiative für Netzwerkinformation E.V. (DINI)
- Digital Repository Infrastructure Vision for European Research (DRIVER)
- La Fundación Española para la Ciencia y la Tecnología (FECYT) i la Red de Bibliotecas Universitarias (REBIUN):  
 José M. Barrueco, et al. "Guía para la evaluación de repositorios institucionales de investigación" (v. 1.0, diciembre 2010).  
<http://www.recolecta.net/buscar/.../GuiaEvaluacionRecolecta1.0-1.pdf>

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Avaluació (3)

- Barrueco, et al., 2010. Visibilitat:
  - Menció del repositori en la pàgina principal de la institució
  - Presència en directoris i recollidors
  - Nom normalitzat
  - URL amigable
  - Iniciatives per fomentar la visibilitat
  - Mínim del 75% en accés obert



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Avaluació (4)

- Barrueco, et al., 2010. Polítiques:
  - Declaració sobre la missió i els objectius del repositori
  - Política pública sobre qui pot dipositar, que i en quins formats
  - Política pública sobre preservació de continguts
  - Política sobre re-ús de metadades
  - Oferta de contacte i assessorament visible



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Avaluació (5)

- Barrueco, et al., 2010. Aspectes legals:
  - L'autor coneix els seus drets
  - L'autor dóna permís per difondre els seus documents
  - Existeix documentació que ajuda a l'autor a decidir
  - Les metadades exportades pel repositori inclouen informació sobre els drets



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Avaluació (6)

- Barrueco, et al., 2010. Metadades:
  - Política d'indexació coneguda pels autors per aconseguir homogenització de les metadades descriptives
  - Sistema de classificació normalitzat
  - Exportació de metadades, a més de DC
  - Format de metadades tècniques i/o preservació



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Avaluació (7)

- Barrueco, et al., 2010. Interoperabilitat:
  - Es proveeixen les metadades a través de l'OAI-PMH
  - Identificació dels recursos en obert en un set DRIVER
  - Es marquen els registres eliminats
  - Testimonis de reanudació d'OAI-PMH de mínim 24h
  - Correu electrònic de l'administrador disponible a *AdminEmail*
  - Camp *Description*
  - Vocabulari DRIVER en el camp *Type*



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Avaluació (8)

- Barrueco, et al., 2010. Logs i estadístiques:
  - Estadístiques sobre l'ús dels documents emmagatzemats

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



### Avaluació (9)

- Barrueco, et al., 2010. Seguretat i autenticitat:
  - Procediment sobre l'elaboració de còpies de seguretat
  - Ús d'identificadors persistents (DOIs, Handles, URNs, etc.)
  - Separació dels continguts per col·leccions

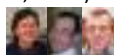
Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



### Avaluació (10)

- Criteris per l'avaluació d'interfícies de consulta:
    - Presència i qualitat de les ajudes
    - Possibilitat de consultes multilingües
    - Modalitats de cerca i llenguatges d'interrogació
    - Dispositius de correcció d'errors en cerques i suggerències a l'usuari
    - Presència i consulta d'índexs
    - Possibilitats de cerca en el text del recurs
- (Sulé, Estivill, Gascón, 2011)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



### Avaluació (11)

- Criteris per l'avaluació d'interfícies de consulta:
    - Capacitat de refinar la cerca
    - Visualització de tesaurus i llistes, i control d'autoritats
    - Presentació dels resultats (ordenació i aplicació de filtres)
    - Serveis de difusió (alertes, etc.)
    - Exportació de referències
- (Sulé, Estivill, Gascón, 2011)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



### Pràctica

- En base als criteris exposats per **visibilitat** (Barrueco et al., 2010) i **interfícies** (Sulé, Gascón, Estivill, 2011), analitza un repositori d'accés obert (disponible a <http://www.openoai.org> o no), determinant:
  1. Dades identificatives del repositori.
    - 1.1. Nom del repositori:
    - 1.2. Institució/ns que el manté/nen:
    - 1.3. Breu descripció del contingut:
    - 1.4. URL d'accés:
  2. Valoració de la seva visibilitat (1 a 3: 1 ha de millorar; 2 suficient; 3 excel·lent).
    - 2.1. Menció del repositori a la pàgina principal de la institució:
    - 2.2. Presència en directoris nacionals i internacionals:
    - 2.3. Presència en recolectors nacionals i internacionals:
    - 2.4. Existència d'un nom normalitzat en tots ells:
    - 2.5. Existència d'una URL amigable:
  3. Valoració de la interfície (1 a 3: 1 ha de millorar; 2 suficient; 3 excel·lent).
    - 3.1. Presència i qualitat de les ajudes:
    - 3.2. Dispositius de correcció d'errors en cerca i suggeriment a l'usuari:
    - 3.3. Capacitat de refinar la cerca:
    - 3.4. Exportació de referències:
  4. Total punts (suma dels anteriors):
  5. Valoració personal del repositori (10 línies màxim).

30' treball individual



Estudi de cas: DDD, el repositori institucional espanyol número 1 del Ranking

### DDD (1)



<http://ddd.uab.cat/>

### Dipòsit Digital de Documents (2)

- La Universitat Autònoma de Barcelona és una universitat pública creada al 1968. Té 3.500 estudiants, 3.500 professors, 83 titulacions de grau, 1.000 de formació permanent i 80 doctorats (2011). El pressupost és de 367 M€ (2010)
- La Biblioteca de la UAB va crear al 2006 el DDD (Dipòsit Digital de Documents), el seu repositori institucional, per recopilar, administrar i difondre els materials produïts per la institució
- Està situat en la primera posició espanyola del Ranking CSIC
- CDS Invenio, 73.000 registres. Majoritàriament de lliure accés (alguns restringits)

UAB



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Dipòsit Digital de Documents (3)

- Objectius del repositori DDD:
  - Difondre els resultats de la investigació universitària
  - Incentivar la creació i ús de la producció científica pròpia
  - Millorar el control bibliogràfic
  - Impulsar l'edició electrònica
  - Augmentar la visibilitat dels documents i dels autors
  - Afegir valor als documents a través de citacions normalitzades, estadístiques de consulta, adreces permanents i mecanismes de preservació



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Dipòsit Digital de Documents (4)



Cristina Azorín és la seva responsable:

“Per crear un repositori cal lluitar contra les llegendes urbanes: no són fàcils de crear, no són econòmics, no són fàcils de mantenir, no sabem si duraran per sempre, i poden acabar amb els usuaris de les biblioteques” (2006)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Dipòsit Digital de Documents (5)

- Està estructurat en 9 col·leccions diferents:

- ☒ **Materials de curs** (36.893)
  - Exàmens (5.502) Guies de l'estudiant (18) Programes de curs (Guies docents) (30.481) Recursos docents (82)
- ☒ **Llibres i col·leccions** (2.113)
  - Llibres (1.848) Capítols de llibres (89) Documents institucionals (51) Guies de Biblioteques UAB (103) Guies de ràdio (22)
- ☒ **Documents de recerca** (6.862)
  - Estudis (541) Patents (73) Prepublicacions (450) Tesis doctorals (2.866) Treballs de recerca i projectes de final de carrera (1.369) Working papers (975)
- ☒ **Publicacions periòdiques** (1.788)
  - Anuaris i memòries (547) Bulletins (1.039) Revistes (202)
- ☒ **Articles** (24.208)
  - Articles publicats (24.208)
- ☒ **Jornades i congressos** (532)
  - Ponències i comunicacions (386) Posters (34) Presentacions (112)
- ☒ **Documents gràfics** (5.594)
  - Cartells (5.287) Fotografies (83) Mapes (83) Pestalls (131)
- ☒ **Multimèdia** (527)
  - Àudio (107) Exposicions de les Biblioteques UAB (321) Vídeos (199)
- ☒ **Fons personals** (10.145)
  - Bernard Lesfargues (336) David Rosenthal (6) Ferran Sunyer i Balaguer (1.159) Jordi Arbonès (946) José Agustín Goytisolo (4.116) Pedro Pascual (1.451) Pere Calders (2.131)

### Dipòsit Digital de Documents (6)

- El repositori s'alimenta per:
  - Autoarxiu
  - Còpia de registres del catàleg de la UAB, més el propi document
  - Recol·lecció de les metadades procedents de repositoris consorciats
  - Càrregues de bases de dades de digitalitzacions retrospectives



“Hem de reconèixer que el creixement dels repositoris respon principalment a la tasca constant del personal de les biblioteques, més que a les polítiques



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Pràctica

- Visita el DDD (<http://ddd.uab.cat/>):
  - Consulta la subcol·lecció d'exàmens (accés restringit).
  - Visita el llistat de treballs de recerca i Fi de carrera dels estudiants.
  - Consulta l'apartat de vídeos de la UAB, produïts per la pròpia Biblioteca.
  - Localitza els documents relatius a Mèxic en el fons personal de l'escriptor Pere Calders.

10' treball individual

## 4. Preservació digital. Panorama

### Panorama (1)

"Año 2054. Mis nietos (que no han nacido) están en el desván de mi casa (que no he comprado). Descubren una carta de 2007 y un CD-ROM. La carta dice que el CD tiene un documento en el que se da la clave para heredar mi fortuna (que no he ganado). Mis nietos sienten curiosidad, pero nunca han visto un disco compacto, salvo en viejas películas. Aun cuando localizaran un lector de disco adecuado, ¿cómo lograrían hacer funcionar los programas necesarios (como Microsoft Word) para interpretar el contenido?"

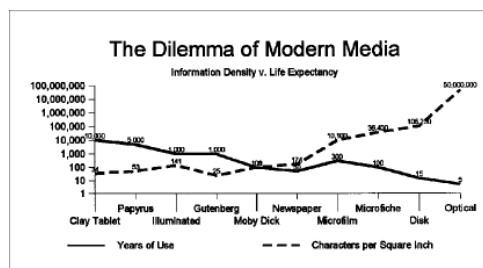
"Si no fuera por la carta, mis nietos no sabrían que deben abrir el CD-ROM. La carta es legible, sin máquinas ni conocimientos a parte del propio idioma. Dado el carácter mutable de programas y circuitería, dentro de 50 años la única información legible será la carta."

Jeff Rothenberg. "¿Son perdurables los documentos digitales?".  
*Investigación y ciencia*, (marzo 1995).  
<http://www.dfists.ua.es/~gil/articulo.pdf>

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013. (Keefer, Gallart, 2007)

### Panorama (2)

Documents primitius: menys informació, més durada.  
Documents digitals: més informació, menys duradera.



(Paul Conway. *Preservation in the digital world*, 1996).  
<http://www.clir.org/pubs/reports/conway2>)

### Panorama (3)

- Anys 80-90: els responsables de preservació tradicional observen amb preocupació l'augment de recursos digitals en els seus fons.

– Saben com preservar documents analògics, però desconeixen què fer amb els nous **suports** (CD-ROM, DVD, accés en línia) i **formats** (text, imatges, hipertext, multimèdia).

"Disponemos de cintas de 1949 que suenan perfectamente, y otras de 1989 que se han perdido completamente"

(Marc Kirkeby, Sony, en relació a l'accés a 600.000 enregistraments musicals en cintes magnètiques)

(Keefer & Gallart, 2007)  
Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.

### Panorama (4)

- Biblioteques, Arxius i Museus no han estat immunes als canvis tecnològics i la irrupció d'Internet:
  - D'una banda, han d'integrar a les seves col·leccions recursos digitals, en formats i suports que no estan consolidats. Han de fer-ho a gran escala. Són els documents **nascuts digitals** (*born digital*).
  - D'altra, han d'aprofitar les oportunitats del medi digital per difondre les seves col·leccions de **documents analògics**, i contribuir també a la seva preservació.

Ciro Llueca. Repositoris digitals. COBDC/BEG 2013.



## Integració dels *born digital*



## Difusió (i preservació) d'anàlegs



## Movimientos Migratorios



<http://pares.mcu.es/MovimientosMigratorios>

## Movimientos Migratorios

- El portal **Movimientos Migratorios Iberoamericanos** coordinat pel **Ministerio de Cultura** amb l'objectiu de facilitar l'accés a la documentació administrativa relativa a la emigració espanyola a Iberoamèrica, i que compta amb la col·laboració de l'**Archivo General de la Nación de México**, l'**Archivo General de la Nación de la República Dominicana**, l'**Archivo General de la Nación de Uruguay**, i la **Fundación Complejo Cultural Parque de España**, d'Argentina, i és prevista la incorporació de l'**Archivo Nacional de la República de Cuba**.
- El sistema inclou registres dipositats directament en el repositori, o bé dipositats en d'altres repositoris associats, sempre que compleixin metadades estàndars.
- Software a mida, 34.000 registres.

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Práctica

- Localitza a Movimientos Migratorios (<http://pares.mcu.es/MovimientosMigratorios>):
  - Persones amb algun dels teus cognoms, o bé:
  - La fitxa del passatger Leoncio Almenara, i identifica els detalls del seu viatge, així com les funcions d'edició de resultats del repositori.

5' treball individual



## Panorama (5)

- "Abarca todas las consideraciones relacionadas con la gestión, administración, financiamiento y recursos humanos que tienden a salvaguardar el bienestar de las colecciones de bibliotecas" (IFLA)
- "Actividades asociadas con el mantenimiento en buenas condiciones del material de archivo y bibliotecario para su utilización en forma original" (ALA)
- "La preservación a largo termini implica la provisión d'accés indefinit als documents o, com a mínim, al contingut d'aquests (...). En el document d'arxiu, la preservació del context i entorn de creació explicarà, complementarà i donarà valor al document" (Soler, 2009)

(Soler, 2009)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013. (Keefer, Gallart, 2007)



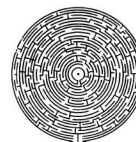
## Panorama (6)

- "Preserving digital information" (Research Libraries Group & Commission on Preservation and Access, USA, 1996)
- "Digital preservation needs and requirements in RLG member institutions" (Hedstrom & Montgomery, 1998)
- "Preservation in the digital world" (Yale University, USA, 1996)
- "Long term preservation of electronic materials" (Joint Information Systems Committee & British Library, UK, 1996)
- CEDARS Digital Preservation Project (Granger et al., 1998-2002)
- CAMALEON (University of Leeds, UK, Rothenberg)
- "PADI: Preserving Access to Digital Information" (National Library of Australia, 1996)
- "Directrices para la preservación del patrimonio digital" (UNESCO, 2003)
- "NEDLIB", "E-Depot" & "DIAS" (Koninklijke Bibliotheek, DK, 1998-2000)
- "Kulturarw3" (Kungliga Biblioteket, SE, 1996)

(Keefer, Gallart, 2007)

## Panorama (7)

- Després de diversos estudis i projectes, no existeix en l'actualitat un sistema informàtic que compleixi tots els requeriments de preservació digital
- Es combinen estratègies a mig plaç i activitats experimentals per garantir l'accés als fons digitals



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Panorama (8)

- Propietatst de la preservació digital:
  - És una **sèrie d'activitats**, no tècniques aïllades.
  - La finalitat és la **utilització futura** dels documents.
  - Es pretén conservar el **document físic, a més de la informació** que conté.

(Keefer & Gallart, 2007)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## 4.1 Estratègia

### Estratègia (1)

- La preservació d'un fons tradicional s'establia en base al material a conservar i al temps de preservació



Les "Homilies d'Organyà" (finals s. XIII) és un dels escrits més antics en català. Descobert el 1904, es conserva a la Biblioteca de Catalunya

(Keefer, Gallart, 2007)

### Estratègia (2)

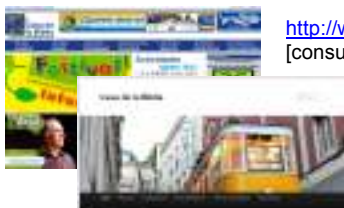
- A diferència de l'actuació tradicional, orientada a mantenir el document en el seu estat original, els documents digitals han de ser modificats i transformats contínuament per a que continuïn sent accessibles i intel·ligibles
- Es tracta de **preservació per alteració**
  - A més de preservar una versió accessible i llegible, els experts recomanen preservar, a més, la cadena de bits original

(Keefer, Gallart, 2007)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Estratègia (3)

- La preservació digital és més complexe:
  - Quins recursos cal preservar?
  - Quan s'ha d'iniciar el tractament?
  - Qui s'implicarà en el procés de preservació?



<http://www.casadelabiblia.org/>  
[consulta febrer 2013]

(Keefer, Gallart, 2007)

### Estratègia (4)

Preservació tradicional (Baró, 1990)	Preservació digital (Jones, Beagrie, 2001)
<ul style="list-style-type: none"> <li>Obligació legal</li> <li>Valor intrínsec del material</li> <li>Pertinença als interessos de la institució</li> <li>Estat físic</li> <li>Ús actual i potencial</li> <li>Valor artístic</li> <li>Polítiques sobre emplaçament i manipulació</li> <li>Disponibilitat en altres biblioteques</li> <li>Cost-Efectivitat del tractament</li> <li>Disponibilitat del personal i instal·lacions</li> <li>Existència de suports alternatius</li> </ul>	<ul style="list-style-type: none"> <li>Les anteriors, i definir si:</li> <li>El material s'ajusta a la política de selecció</li> <li>Una altra institució ha assumit la responsabilitat</li> <li>Hi ha obligació legal</li> <li>El valor del document justifica la inversió</li> <li>La institució pot manipular el document</li> <li>El recurs és llegible avui i en el futur</li> <li>El recurs està en un suport que serveix per transferir o emmagatzemar les dades</li> <li>El recurs disposa de documentació identificativa, incloent metadades</li> </ul>

(Keefer, Gallart, 2007)

### Estratègia (5)

- Tot i existint capacitat tècnica i voluntat institucional de preservació digital, els recursos limitats obliguen a les institucions a establir una política:
  - Neccessitats i recursos de la institució
  - Interessos dels usuaris
  - Diferents aspectes físics i intel·lectuals

(Keefer, Gallart, 2007)

### Estratègia (6)

- L'estratègia requereix trobar l'equilibri entre la importància relativa del document i els factors de risc



(Keefer, Gallart, 2007)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Estratègia (7)

- Actuació reactiva o proactiva?
  - La solució reactiva no és recomanable per objectes digitals: al detectar l'error, sol ser massa tard
  - El "llarg termini" en objectes digitals és de 20 anys (Waller, Sharpe, 2006)



1986



1987



1996



2006

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013 (Keefer, Gallart, 2007)

### Estratègia (8)

- Actuació esporàdica o continuada?
  - La solució esporàdica permet allargar la vida dels documents analògics, però difícilment dels digitals
  - Les intervencions digitals afecten al document i a la seva descripció
  - La intervenció continuada requereix financiació permanent



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013 (Keefer, Gallart, 2007)

### Estratègia (9)

- Actuació o omisió?
  - L'abandonament benigne (*benign neglect*) servia en alguns casos per salvar la vida del document
  - En l'entorn digital, l'oblit és letal. Ha de ser llegible el format i també el suport



#### Sinopsis

Material fotográfico inédito de los insignes fotógrafos Robert Capa, Gerda Taro y David Seymour aparecieron en varias cajas de un apartamento de México D.F. Esta magnífica serie de negativos representan una extraordinaria colección de imágenes que representan diferentes escenas de la contienda civil española (1936-1939) en el frente de batalla, donde se capta perfectamente el horror, el caos de la guerra, de cómo el fascismo y el comunismo se enfrentan en el preludio de la II Guerra Mundial.

WordStar is highly flexible and very visible. Watch the screens as you give commands, and information in various parts of the screen will guide you. You won't see all the information all the time, but it will be there when you need it.

Wordstar 1992

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013. (Keefer, Gallart, 2007)

### Estratègia (10)

- Actuació al crear el document o en cas de deteriorament?
  - És recomanable, tot i els reptes (creació de documents descentralitzada, inèrcia dels autors i departaments) establir pautes en la creació, i assessorar (selecció de formats, suports, metadades)



(Keefer, Gallart, 2007)

### Estratègia (11)

- Hem de decidir:
  - Preservar el contingut, inclús millorar la llegibilitat o utilitat mitjançant les tecnologies disponibles en el moment de la seva recuperació
  - Mantenir el recurs mitjançant la preservació de l'aspecte i comportament original, amb les limitacions que poden representar



(Keefer, Gallart, 2007)

### Estratègia (12)

- Valors essencials a preservar en documents d'arxiu:
  - Fiabilitat: és possible conèixer la traçabilitat, des de que fou creat per algú, fins el present.
  - Autenticitat: deixant clar què és el màster i què còpies de difusió, ús, etc.
  - Integritat: ha de restar intacte, tal com mostrin les seves metadades.
  - Identitat: es coneixen totes les persones que han creat o treballat en el document.
  - Capacitat d'ús i accés per a l'ull humà.

(Soler, 2009)

(Elvira, 2011)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Estratègia (13)

- Prioritzant:
  - Documents nascuts digitals
  - Documents digitalitzats sense versió analògica
  - Documents digitalitzats amb versió analògica

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Pràctica

- Defineix una política de preservació per a la teua unitat d'informació / institució:
  - Quina documentació es genera? (Classifica-la en grans grups)
  - Què es preserva i per què?
  - Com afecta la preservació als processos diaris?

10' treball individual  
+ 10' presentació



## 4.2 Tàctica

### Tàctica (1)

- Existeixen dues accions bàsiques de preservació digital, que habitualment actuen combinades:
  - Actualització dels suports
  - Migració dels formats



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Tàctica (2)

- Actualització / *Refreshing*
  - Transferència de dades d'un suport a un altre
  - No preserva les dades, però és un pas repetible i imprescindible per garantir l'accés a aquestes
  - Quina previsió de vida té el nou suport?

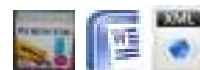


(Keefer, Gallart, 2007)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Tàctica (3)

- Migració
  - Transferència de dades d'un **format** a un altre
  - L'estratègia més usada, tot i existir el risc de perdre elements del document
  - És una operació experimentada, sense coneixements tècnics especialitzats
  - És difícil programar la migració en objectes complexos, altera els documents, i s'ha de repetir al llarg del temps



(Keefer, Gallart, 2007)

### Tàctica (4)

- Emulació
  - **Simular** el comportament del programari original
  - L'estratègia més respectuosa amb el document original
  - No requereix un seguiment continuat del format, soluciona la lectura per objectes complexos
  - Existeixen poques experiències reals, els emuladors han d'actualitzar-se, requereix coneixements especialitzats ara i en el futur



(Keefer, Gallart, 2007)

### Tàctica (5)

- Accions complementàries:
  - **Museu tecnològic**: conservar tots els models, còpies i peces de reparació, per poder emular
  - **Replicació**: preservar múltiples còpies digitals en línia, en formats estàndards (Castillo, Jorba, 2008)
  - **Analogització**: apostar per suports de solvència contrastada en preservació: paper, microformes, etc.
  - **Arqueologia digital**: confiar en el futur



(Keefer, Gallart, 2007; Serra 2009)

## Tàctica (6)

- Riscos per la longevitat:
  - Visualització: en el futur hauran de conèixer la infraestructura original
  - Interferències: a causa d'encriptació o compressió, la recuperació pot ser problemàtica
  - Interrelacions: els enllaços d'hipertext es poden perdre
  - Custòdia: qui es responsabilitza?
  - Traducció: salvar el *look & feel*



(Besser, citat per Keefer, Gallart,

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013. 2007)

## CLAMOR



<http://digital.march.es/clamor/>

## CLAMOR

- La Fundació Juan March creada al 1955 pel financer espanyol Juan March Ordinas. És una institució privada que desenvolupa les seves activitats en el camp de la cultura humanística i científica. Té seu a Madrid, y és titular del Museo de Arte Abstracto Español (Cuenca) i del Museo Fundación Juan March (Palma de Mallorca)
- Al 2011 ha obert el seu repositori digital: CLAMOR - Colección Digital de Música Española, amb l'objectiu de recopilar, preservar i donar accés a la activitat concertística de la Fundació, des de 1975
- Conté gravacions en viu, programes de mà, informació sobre els intèrprets, fotografies, etc.
- Fedora (adaptació Islandora)



## Pràctica

- Visita **CLAMOR** (<http://digital.march.es/clamor/>):
  - Comprova el sistema de funcionament i visualització de resultats per cerca i per navegació.
  - Accedeix a la "suite Iberia" d'Isaac Albéniz i comprova la documentació disponible.

5' treball individual



## El suport (1)

- Factors de degradació:
  - La composició "en capes" dels suports, en les quals s'energistren (un cop i un altre) les dades per mitjà d'un procés magnètic poden causar envelliment precoç, que pot provocar errors en la lectura de les dades.

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013. (Keefer, 2012)

## El suport (2)

- Factors de degradació:
  - Temperatura i humitat
  - Baixa qualitat dels materials originals
  - Baixa qualitat del procés de fabricació
  - Falta de controls ambientals en la fabricació, transport i emmagatzegament
  - Proximitat a camps magnètics
  - Brutícia, pols, en l'entorn
  - Introducció de bacteris, fongs
  - Mal estat dels equips d'enregistrament i de lectura

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013. (Keefer, 2012)

### El suport (3)

- Alguns factors de degradació són controlables per la institució:
  - Flexibilitzant els criteris de compra, que dificulten l'adquisició de material d'alta qualitat
  - Promovent el canvi en la inèrcia o reticència de les maneres de treballar
  - Facilitant la capacitat tècnica o econòmica per efectuar canvis en l'entorn físic

(Keefer, 2012)

### El suport (4)

Esperança de vida segons condicions ambientals (humitat relativa i temperatura):

Suport	25 h.r. / 10 C	40 h.r. / 20 C	50 h.r. / 28 C
D3 cinta magnètica	50 anys	15 anys	1 anys
DLT cinta magnètica	75 anys	15 anys	1 anys
CD/DVD	75 anys	20 anys	2 anys
CD-ROM	30 anys	3 anys	3 anys

(Síntesi de Jones & Beagrie, 2001, p. 130)

### El format (1)

- L'elecció d'un format pot condicionar:
  - La selecció o no del recurs per ser preservat.
  - El mètode usat per la seva preservació.
  - La facilitat de recuperar i visualitzar el document.
  - El tipus d'ús possible en el futur.
  - El cost de la preservació.

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013. (Keefer, 2012)

### El format (2)

- La importància d'escollir un format:
  - Format estàndard, reconeixible per una àmplia gama de programes (.txt)
  - Format propietari amb gran base de clients, que puguin arribar a incentivar econòmicament les solucions (.doc)
  - Format propietari de difusió limitada (.dxf)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013. (Keefer, Gallart, 2007)

### El format (3)

- Existeixen ca. 16.000 formats (NARA)
  - Document (doc, rtf, pdf, xsd, ps)
  - Imatge (jpeg, gif, png, psd, tif)
  - Audio (wav, mp3, aac)
  - Video (wmv, avi, rm, mpg)
  - Full de càlcul (xls, xsc)
  - Estadística (SPSS)
  - Diagrama o disseny (Visio, AutoCAD)
  - Bases de dades (SQL, MySQL, Oracle, Access)
  - Presentacions (ppt)
  - Pàgines web
  - Etc.



<http://www.gdfr.ir>

(Bates et al., 2006, citat per Keefer, Gallart, 2007)

### El format (4)

Formats	Preferit	Acceptable	Problemàtic
Dades estadístiques	SPSS, por, sav	STATA, xport, text separat per tabulació	Text d'amplada fixe (no delimitat)
Base de dades	SQL separat per tabuladors o comes	ACCESS, dBase, FilemakerPro, FoxPro, Paradox	Text d'amplada fixe (no delimitat)
Full de càlcul	Text delimitat separat per tabuladors	Excel, Lotus, Quattro Pro	
Text	Enriquit, rtf, xml, txt	MsWord, HTML, Wordperfect, paper	CAQDAS, WinMax, Nudist
Imatge	tiff, png	gif, bmp	Adobe PDF
Video	MPEG	Quicktime	
So	mp3	wav	

(Recomanacions UK data Archive, citat per Keefer)

### Pràctica

- Identifica, per mitjà del projecte **PRONOM** (<http://www.nationalarchives.gov.uk/PRONOM>), el format dels següents documents, trobats en un vell disc dur:
  - El fitxer testament.wp5
  - Un arxiu anomenat bestofmarysempere.ogg
  - Un arxiu molt pesat anomenat whokilledkennedy.lzw
- Identifica el format més convenient per la seva preservació.

10' treball individual



### Preservem digitalitzant?(1)

- La digitalització s'usa com a mètode de preservació (i difusió) dels documents analògics, atès que pot promoure:
  - Cerca i opcions de localització més fàcils.
  - Accés sense barreres temporals, físiques o geogràfiques.
  - Més maneres de visualitzar el document.
  - Millor qualitat de la imatge i el text.
  - Fàcil i ràpida reproducció, sense agressions a l'original.
  - Nous usos, com l'anàlisi d'imatge i text, comparació de documents allunyats, etc.

(Keefer, 2012)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Preservem digitalitzant?(2)

- Però lògicament, la digitalització no substitueix als processos preservació en els documents analògics:
  - Baixa manipulació.
  - Temperatura i humitat constants.
  - Condicions lumíniques.
  - Insectes, desastres naturals, robatoris o pèrdues, etc.

(Keefer, 2012)



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Preservem digitalitzant?(3)

- Prioritzem?
  - Documents fràgils
  - Singularitat
  - Interés per a la recerca
  - Requeriments d'ús
  - Completar col·leccions
  - Freqüència d'ús i de consulta
  - Domini públic

(Serra, 2011)



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Preservem digitalitzant?(4)

- Preparar l'escàner i els documents:
  - Neteja de les lents i de l'entorn de l'escàner
  - Neteja i ordre dels documents.
- Ajustar els aspectes tècnics:
  - Blanc i negre: mecanuscripts, documents impresos.
  - Escala de grisos: manuscrits, música impresa.
  - Color: mapes, fotos, il·lustracions.


[www.digitalpreservation.gov](http://www.digitalpreservation.gov)

### Preservem digitalitzant?(5)

Tipus documento	Preservació	Accés
Text BN/grisos (llibres, revistes)	imprès TIFF 300 dpi	JPEG, PDF 150 dpi
Text imprès color (manuscrits, llibres il·lustrats)	TIFF 300 dpi	JPEG, PDF 150 dpi
Fotografies, mapes	TIFF o JPEG2000 500 dpi	JPEG, PDF 300 dpi
Sonors	WAVE o BWF	MP3
Audiovisuals	MPEG2 o AVI	MPEG2

Més concrecions a:

<http://www.library.cornell.edu/ims/image%20deposit%20guidelines.pdf>

## Preservem digitalitzant?(6)

- Accés i emmagatzematge:
  - Carpetes ben organitzades, estructura arbòria, noms i dates clares.
  - Un mínim de 2 còpies en dos suports diferents, de qualitat fiable, allunyats entre si.
  - Etiquetatge curós (mínims: impressió de pantalla per obtenir un sumari, que acompanya les còpies externes al PC.)
  - Emmagatzematge segur.
  - Noves còpies cada 5 anys, com a mínim



[www.digitalpreservation.gov](http://www.digitalpreservation.gov)

## L'arxiu de la Web



## Pràctica

- Visita Internet Archive (<http://www.archive.org/>):
  - Localitza versions antigues de les pàgines web que vulguis.
  - Comprova el sistema de funcionament i accés als documents.
- Visita PANDORA (<http://pandora.nla.gov.au/>):
  - Identifica una URL d'un tema del teu interès.
  - Compara el sistema de funcionament amb Internet Archive.

10' treball individual



## Metadades (1)

- Sense metadades, el document “preservat” digitalment pot ser il·legible.
  - Metadades generades automàticament.
  - Metadades generades pels creadors.
  - Metadades generats pel personal de la biblioteca, arxiu, etc.



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Metadades (2)

- Sense les metadades no podem garantir que un recurs sigui recuperable i usable:
  - Sense dades descriptives, no es podrà localitzar.
  - Sense dades tècniques, no es podrà visualitzar o utilitzar.
  - Sense dades administratives, no es tindrà la garantia de la integritat ni autenticitat.
  - Sense dades legals, els responsables de la custòdia no sabran quins usos són admissibles o no.

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Metadades (3)

- Dades sobre dades
- Metadada és aquella informació descriptiva sobre el context, qualitat, condició o característica d'un recurs documental, dada o objecte, amb l'objectiu de facilitar la seva recuperació, autenticació, avaluació, preservació i/o interoperabilitat

(Wikipedia)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



### Metadades (4)

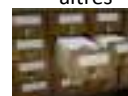
- Els documents que s'introdueixen en el repositori han d'estar descrits usant un conjunt de metadades basades en estàndards internacionals per a facilitar el seu ús posterior
- Han d'estar mínimament normalitzades
- Són, en part, l'equivalent a la descripció bibliogràfica en un catàleg clàssic



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Metadades (5)

- Tipus de metadades
  - Descriptives: creador, títol, data de publicació, tipus de document, etc.
  - Administratives: propietat, drets i usos, equivalent analògic, etc.
  - Tècniques i de preservació: format, resolució, compressió, dispositiu de captura, estructura del document, historial de canvis, etc.
  - Estructurals: relacions amb altres objectes



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Metadades (6)

- Són sistemes estàndards de metadades:
  - DC (Dublin Core)
  - EAD (Encoded Archival Description) > arxius
  - MARC 21 > biblioteques / catàlegs
  - METS (Metadata Encoding and Transmission Standard) > tècniques, preservació
  - PREMIS (Preservation Metadata: Implementation Strategies) > preservació, traçabilitat
  - MODS (Metadata Object Description Schema)
  - TEI (Text Encoding Initiative)

<ead>

TEI

PREMIS MODS



### Metadades (7)

- El **Dublin Core** (1995) és l'estàndard més usat, especialment la seva versió **qualificada**
- Pràcticament tots els programaris de repositoris l'inclouen o el suporten
- Consta de 15 elements descriptius:
  - Títol, creador, matèries i paraules clau, descripció, editor, contribucions, data, tipus de recurs, format, identificador, font, llengua, cobertura, gestió dels drets



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Metadades (6)

- L'elecció del sistema de metadades és essencial per a la seva futura interoperabilitat i visibilitat, més enllà de la nostra pròpia institució
- El sistema d'autoarxiu pot tenir conseqüències en la correcta assignació de metadades
- Habitualment en l'autoarxiu s'opta per uns mínims (autor, títol, data, etc.), que després completaran els serveis tècnics



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Metadades (7)

- La majoria de repositoris compten amb el protocol d'interoperabilitat creat per la Open Archives Initiative, el Protocol for Metadata Harvesting (OAI-PMH)
- El protocol serveix per capturar (si el repositori és agregador) o proveir (si és proveïdor) metadades de/a altres repositoris

<http://www.openarchives.org/pmh/>



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Metadades (8)

- Si en el repositori es barregen diferents tipus de col·leccions (científica, docent, cultural, institucional, etc.) és aconsellable diferenciar-les:
  - Per intercanvis futurs
  - Per participar amb cada col·lecció en diferents repositoris



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Metadades (9)

- Per participar en portals/agregadors és necessari complir determinades especificacions
  - A DRIVER només documents de recerca
  - A OpenAIRE només documents finançats publicament pel 7PM
- Aquests portals/agregadors es limiten a la integració i recuperació de metadades: el text complet sempre està en el repositori d'origen (cobertura legal)



Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

D

<http://digital.csic.es>



Isabel Bernal, és la responsable:

- “Aspiramos a ser la memoria digital de la ciencia española”
- “El futuro: El repositorio substituirá al catálogo; automatización de los procesos de publicación de citas para evitar duplicaciones; cambio de chip: las bibliotecas orientadas a favorecer la difusión de la producción.... ¿Cuál es el límite de los softwares? ¿Podrán convertirse en bibliotecas digitales, más allá de ser almacenes?”

### Digital CSIC

- El CSIC (Centro Superior de Investigaciones Científicas) és la major agència pública espanyola d'investigació científica (3a UE), amb 15.000 persones en plantilla (d'aquestes 3.000 són investigadors), genera el 20% de la producció científica espanyola
- Va signar la Declaració de Berlín al 2006
- El repositori Digital CSIC (2008-) facilita l'accés a articles (especialment *postprint*), llibres, literatura gris, patents, bases de dades, comunicacions a congressos, material didàctic i divulgatiu, informes tècnics, etc.
- DSpace, 38.000 registres, 180.000 visites mensuals, #2 ranking repositorios (.es)
- Els incentius per als investigadors: augmentar la visibilitat de la seva producció científica; preservar tota la producció; tractament centralitzat (40% oficina tècnica + 40% biblioteques + 15-20% autoarxiu)

### Pràctica

- Busca un registre del teu interès en el **Digital CSIC** (<http://digital.csic.es/>).
- Compara la fitxa descriptiva que es troba per defecte amb l'estructura **Dublin Core** opcional.
- Visita lliurament el repositori **Digital CSIC**.

15' treball individual



### Repositori segur (1)

- La majoria de repositoris institucionals no poden ser considerats com repositoris de preservació
- Un repositori de preservació té com a responsabilitat única dur a terme la preservació a llarg plaç dels seus objectes digitals (PREMIS 2008)

(Termens, 2009)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## Repositori segur (2)

- Establir un repositori segur:
  - Suports:
    - Emmagatzemar els màsters en suports digitals d'alta qualitat
    - Comprovar la llegibilitat dels suports periòdicament
  - Còpies de seguretat:
    - Copiar els màsters i dipositar en lloc extern segur
  - Dispositius d'enregistrament i de lectura
    - Seguiment per garantir que siguin de bona qualitat, i estiguin en bon estat

(Establishing a central depository..., p. 16-17  
<http://www.library.cornell.edu/ims/image%20deposit%20guidelines.pdf>)

## Repositori segur (3)

- Establir un repositori segur:
  - Emmagatzegament:
    - Conservar en entorn controlat de temperatura i humitat
    - Mantenir nivells constants més que seguiment estricte de nivells òptims
    - Allunyat de camps magnètics
    - Mantenir net l'entorn d'operacions
    - Minimitzar la manipulació i l'ús dels suports
  - Seguretat:
    - Control d'accés
    - Accés tan sols a l'entorn de treball

(Establishing a central depository..., p. 16-17  
<http://www.library.cornell.edu/ims/image%20deposit%20guidelines.pdf>)

## Repositori segur (4)

- Establir un repositori segur:
  - Refresh:
    - Passe periòdic a nous suports (en comprar nous dispositius d'emmagatzegament i lectura?)
    - Seguir un procés de verificació (checksum) per garantir autenticitat i integritat posterior al refresh
  - Documentació
    - Explicar per escrit tota acció duta a terme durant el procés de manteniment o refresh

(Establishing a central depository..., p. 16-17  
<http://www.library.cornell.edu/ims/image%20deposit%20guidelines.pdf>)

## Repositori segur (5)

- El model OAIS (Open Archival Information System) va ser desenvolupat per la NASA, i es va convertir en la norma ISO 14721:2012
- Serveix per referir-se a conceptes i processos en base a una norma compartida
- No ofereix un disseny específic d'arquitectura, és una fulla de ruta



"Vehículo de cuatro ruedas, capaz de transportar entre dos y ocho personas con una velocidad posible de 130 km/h y un espacio para equipaje y con prestaciones para que el viaje sea más agradable (radio, aire climatizado, asientos cómodos...)" (Gladney 2006)

(Keefer, Gallart)

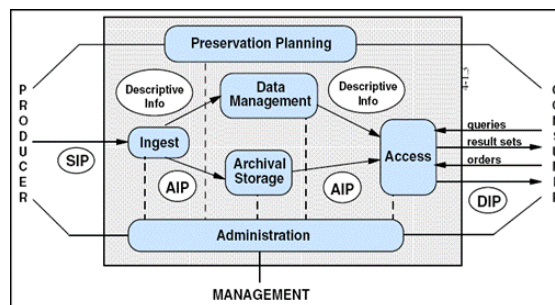
## Repositori segur (6)

- El model OAIS esbossa una arquitectura per al paquet d'informació (objecte+metadades):
  - Ingest: incorporació al repositori
  - Archival storage: emmagatzemament a llarg plaç
  - Data management: la seva gestió
  - Access/dissemation: el seu accés i difusió
  - Preservation planning: preservació planificada
  - Archive administration: gestió del repositori

(Keefer, Gallart, 2007)

Ciro Ulluca, Repositoris digitals, COBDC/BEG 2013.

## Repositori segur (7)



Ciro Ulluca, Repositoris digitals, COBDC/BEG 2013.

### Repositori segur (8)

- Les institucions han de poder ser fiables:
  - Garantitzar la integritat i autenticitat del recurs
  - Vetllar per la protecció dels drets
  - Limitar l'accés a usuaris autoritzats
  - Mantenir les eines necessàries per identificar i recuperar la informació
  - Assegurar la presència de metadades normalitzades

"Trusted digital repositories: attributes and responsibilities" (RLG & OCLC, 2002)

(Keefer, Gallart, 2007)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Repositori segur (9)

- Els repositoris es poden certificar:
  - Per exemple, en planificació i estratègies de preservació, el repositori haurà de:
    - Documentar la seva estratègia de preservació
    - Realitzar estratègies per emmagatzemar i migrar
    - Usar informació tècnica disponible internacionalment
    - Tenir un mecanisme per detectar quan un format està en risc d'obsolescència
    - Etc.

"Audit checklist for the certification of trusted digital repositories" (RLG & NARA, 2005)

(Keefer, Gallart, 2007)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### Repositori segur (10)

- Els repositoris es poden certificar:
  - Organització
  - Gestió de l'objecte
  - Infraestructura i seguretat

"Catalogue of criteria for trusted digital repositories" (NESTOR, Network of Expertise in Long-Term Storage of Digital Resources, 2006)

(Keefer, Gallart, 2007)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



### Repositori segur (11)

- El repositori serà més segur:
  - Si disposa d'un sistema de rèplica en altres institucions
  - Si disposa d'una caixa forta o arxiu fosc, d'accés restringit



### Repositori segur (12)

- Externalització dels serveis de preservació:
  - El col·lectiu bibliotecari accepta com un fet impossible que una sola institució pugui afrontar totes les exigències de la preservació digital
  - Més enllà de confiar amb la fiabilitat a llarg plaç dels serveis comercials, s'ha fomentat la creació de centres sense ànim de lucre:
    - Digital Preservation, California Digital Library
    - European Archive
    - UK Data Archive
    - Metarchive.org

(Keefer, Gallart, 2007)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.



### CBUC

- Al 2007 va encarregar un estudi a A. Keefer i N. Gallart per garantir la preservació dels documents dipositats en els seus repositoris
- L'informe (en català està disponible a RECERCAT) finalitza amb una sèrie de recomanacions pel CBUC i les institucions participants:
  1. Determinar per a cada repositori quin nivell de preservació es vol garantir
  2. Acordar i documentar les pràctiques sobre obtenció d'autoritacions
  3. Acordar i documentar les pràctiques per comprovar la integritat de les obres
  4. Establir recomanacions per crear paquets (objectes + metadades administratives i tècniques)
  5. Decidir un calendari per implementar l'emmagatzematge d'arxiu



## CBUC

6. Col·laborar amb un mínim de 2 institucions per disposar d'una xarxa de rèpliques anti-desastre, i d'un arxiu fosc
7. Sol·licitar al soci tecnològic que se sotmeti a certificació i que apliqui la normativa
8. El·laborar i publicar la política de preservació
9. Cooperar amb alguna xarxa internacional de prestació de serveis per garantir la preservació
10. Definir una estratègia de vigilància tecnològica per controlar l'evolució de la tecnologia



## CBUC

- El Consorci de Biblioteques Universitàries de Catalunya (CBUC) va ser creat al 1996 com organisme públic autònom per millorar els serveis bibliotecaris a través de la cooperació. Els seus membres són 14 universitats catalanes i quatre organismes d'investigació
- Entre els seus projectes: CCUC, TDX (el primer repositori a Espanya) i d'altres repositoris patrimonials i de producció científica



**Lluís Anglada** és el seu director: "El naixement del CBUC va ser fruit de diverses coses, de la oportunitat de l'amistat, del moment històric, però sobretot és fill de la voluntat de fer (...) de tenir present que per millorar no n'hi ha prou amb millorar només a casa nostra, sinó que cal millorar tot l'edifici en el que vivim i si podem també el carrer"



## MetaArchive: què és?

Una xarxa d'institucions que tenen interès en preservar d'una manera efectiva i amb baix cost alguns dels seus conjunts de dades.

Per fer-ho proposen fer còpies de les dades creuades entre les seves institucions.



## MetaArchive: Qui en forma part?

Auburn University  
Boston College  
Clemson University  
Florida State University  
Folger Shakespeare Library  
Georgia Tech  
Indiana State University  
Library of Congress  
Penn State University  
Pontifícia Universidade Católica do Rio de Janeiro (Brasil)  
Rice University  
University of Hull (Anglaterra)  
University of Louisville  
University of North Texas  
University of South Carolina  
Virginia Tech



## MetaArchive: Funcionament tècnic

- Usa el programari LOCKSS (Lots of Copies Keep Stuff Safe). Programari molt utilitzat des de fa anys i que ha demostrat la seva efectivitat.
- En un servidor nosaltres guardariem (no en obert) tesis d'altres universitats i ells les nostres.
- En cas de **pèrdua de fitxers** el sistema les recupera de la xarxa.
- En cas de **corrupció de fitxers** el sistema també els recupera de la xarxa (checksums).



## COFRE



## 1. Què són els repositoris digitals i per a què serveixen (síntesi)

### En síntesi (1)

- L'accés obert és un moviment que promou l'accés universal al coneixement finançat amb diners públics
- Va sorgir a Internet com a resposta a una sèrie de circumstàncies econòmiques i culturals que han afectat a les biblioteques, arxius i museus

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### En síntesi (2)

- Els repositoris digitals són “un conjunto de servicios prestados por una institución a su comunidad para recopilar, gestionar, difundir y preservar su producción digital a través de una colección organizada, de acceso abierto e interoperable”

(Barrueco et al., 2010)

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### En síntesi (3)

- Serveixen per publicar a Internet, normalment amb accés lliure, una base de dades que inclou els documents digitals o digitalitzats que formen els fons i col·leccions relacionades amb una unitat d'informació
- Compartir i preservar!

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### En síntesi (4)

- Existeixen repositoris institucionals, cooperatius, temàtics i portals
- Existeixen passarel·les i sistemes que ofereixen accés a múltiples repositoris

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## 2. Disseny i implementació d'un nou repositori (síntesi)

### En síntesi (1)

- Les unitats d'informació han de conèixer la planificació estratègica de la institució (i participar, si és possible)
- Iniciar un procés de disseny i implementació d'un repositori digital ha de partir necessàriament d'un pla que sigui coherent amb l'entitat que finançarà i impulsarà la nova eina

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### En síntesi (2)

- Les lleis de propietat intel·lectual vigents en la majoria de països obliguen a obtenir permisos dels propietaris dels drets morals i/o materials per publicar en els repositoris, i en quins termes
- Els administradors de repositoris han de tenir el permís dels autors o propietaris dels drets per difondre les obres incloses

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### En síntesi (3)

- L'elecció entre un sistema local o extern de maquinari depèn de diversos factors, bàsicament econòmics
- Si la institució no té un equip informàtic estable i personal tècnic, serà més ràpid i econòmic començar ubicant el repositori en un sistema extern
- A llarg plaç, però, serà més car mantenir una estructura externa

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

### En síntesi (4)

- L'elecció del programari es realitza en base al tipus de document a incloure i el volum previst d'usuaris
- No existeix un únic programari amb solucions per tots els casos
- DSpace, un programari lliure, és el més usat

Ciro Lluca. Repositoris digitals. COBDC/BEG 2013.

## 3. Difusió i avaluació (síntesi)

### En síntesi (1)

- En paral·lel a comptar amb cobertura estratègica i legal, i a definir les necessitats de maquinari i programari, és necessari:
  - Definir les col·leccions que formaran part del repositori (pensant en clau d'impacte)
  - Identificar els recursos que ens calen per a la implementació
  - Determinar el sistema de metadades, essencial per a la seva futura interoperabilitat i visibilitat



### En síntesi (2)

- La majoria de repositoris compten amb el protocol d'interoperabilitat creat per la Open Archives Initiative, el Protocol for Metadata Harvesting (OAI-PMH)
- Aquest protocol serveix per capturar (si el repositori és agregador) o proveir (si és proveïdor) metadades de/a altres repositoris

### En síntesi (3)

- Les estratègies de comunicació han de considerar als dipositants potencials i també als usuaris
- Establir aliances és imprescindible: directius, departaments paral·lels, unitats d'informació, altres repositoris, etc.
- Ús de la web social (amb un pla d'acció)

### En síntesi (4)

- Existeixen diverses guies per avaluar els repositoris
- Són eines a considerar abans, durant, i després de la implementació
- Alguns dels aspectes a avaluar: la visibilitat, accés a les polítiques del repositori, aspectes legals, interoperabilitat, etc.

## 4. Preservació digital (síntesi)

### En síntesi (1)

- Després de diversos estudis i projectes, no existeix en l'actualitat un sistema informàtic que compleixi tots els requeriments de preservació digital
- Es combinen estratègies a mig plaç i activitats experimentals per garantir l'accés als fons digitals

### En síntesi (2)

- A diferència de l'actuació tradicional, orientada a mantenir el document en el seu estat original, els documents digitals han de ser modificats i transformats contínuament al llarg del temps per a que continuïn accessibles i intel·ligibles
- Es tracta de la preservació per alteració
  - A més de preservar una versió accessible i llegible, els experts recomanen preservar, a més, la cadena de bits original

(Keefer, Gallart,  
2007)

### En síntesi (3)

- Existeixen dues accions bàsiques de preservació digital, que habitualment actuen combinades:
  - Actualització dels suports
  - Migració dels formats

### En síntesi (4)

- El model OAIS esbossa una arquitectura per al paquet d'informació (objecte+metadades):
  - Ingest: incorporació al repositori
  - Archival storage: emmagatzemament a llarg plaç
  - Data management: la seva gestió
  - Access/dissemination: el seu accés i difusió
  - Preservation planning: preservació planificada
  - Archive administration: gestió del repositori

(Keefer, Gallart, 2000)

Gràcies!  
Estic a la vostra disposició!

Ciro Llueca, [cllueca@gmail.com](mailto:cllueca@gmail.com)